```
   1 CGGGCAGCAA AGGAGGATGG CGAGGGGCTG ATACTGAACC CGGGAAGGGT
  51 GGGCTGTGCT GAAGCTAGAG CCGGAGCCGG AGCTGGGGCC AGAACCCGAG
 101 CACTGCCATG TCCACGCAGA GACTTCGGAA TGAAGACTAC CACGACTACA
 151 GCTCCACGGA CGTGAGCCCT GAGGAGAGCC CGTCGGAAGG CCTCAACAAC
 201 CTCTCCTCCC CGGGCTCCTA CCAGCGCTTT GGTCAAAGCA ATAGCACAAC
 251 ATGGTTCCAG ACCTTGATCC ACCTGTTAAA AGGCAACATT GGCACAGGAC
 301 TCCTGGGACT CCCTCTGGCG GTGAAAAATG CAGGCATCGT GATGGGTCCC
 351 ATCAGCCTGC TGATCATAGG CATCGTGGCC GTGCACTGCA TGGGTATCCT
 401 GGTGAAATGT GCTCACCACT TCTGCCGCAG GCTGAATAAA TCCTTTGTGG
 451 ATTATGGTGA TACTGTGATG TATGGACTAG AATCCAGCCC CTGCTCCTGG
 501 CTCCGGAACC ACGCACACTG GGGAAGACGT GTTGTGGACT TCTTCCTGAT
 551 TGTCACCCAG CTGGGATTCT GCTGTGTCTA TTTTGTGTTT CTGGCTGACA
 601 ACTTTAAACA GGTGATAGAA GCGGCCAATG GGACCACCAA TAACTGCCAC
 651 AACAATGAGA CGGTGATTCT GACGCCTACC ATGGACTCGC GACTCTACAT
 701 GCTCTCCTTC CTGCCCTTCC TGGTGCTGCT GGTTTTCATC AGGAACCTCC
 751 GAGCCCTGTC CATCTTCTCC CTGTTGGCCA ACATCACTAT GCTGGTCAGC
 801 TTGGTCATGA TCTACCAGTT CATTGTTCAG AGGATCCCAG ACCCCAGCCA
 851 CCTCCCCTTG GTGGCCCCTT GGAAGACCTA CCCTCTCTTC TTTGGCACAG
 901 CGATTTTTTC ATTTGAAGGC ATTGGAATGG TTCTGCCCCT GGAAAACAAA
 951 ATGAAGGATC CTCGGAAGTT CCCACTCATC CTGTACCTGG GCATGGTCAT
1001 CGTCACCATC CTCTACATCA GCCTGGGGTG TCTGGGGTAC CTGCAATTTG
1051 GAGCTAATAT CCAAGGCAGC ATAACCCTCA ACCTGCCCAA CTGCTGGTTG
1101 TACCAGTCAG TTAAGCTGCT GTACTCCATC GGGATCTTTT TCACCTACGC
1151 ACTCCAGTTC TACGTCCCGG CTGAGATCAT CATCCCCTTC TTTGTGTCCC
1201 GAGCGCCCGA GCACTGTGAG TTAGTGGTGG ACCTGTTTGT GCGCACAGTG
1251 CTGGTCTGCC TGACATGCAT CTTGGCCATC CTCATCCCCC GCCTGGACCT
1301 GGTCATCTCC CTGGTGGGCT CCGTGAGCAG CAGCGCCCTG GCCCTCATCA
1351 TCCCACCGCT CCTGGAGGTC ACCACCTTCT ACTCAGAGGG CATGAGCCCC
1401 CTCACCATCT TTAAGGACGC CCTGATCAGC ATCCTGGGCT TCGTGGGCTT
1451 TGTGGTGGGG ACCTATGAGG CTCTCTATGA GCTGATCCAG CCAAGCAATG
1501 CTCCCATCTT CATCAATTCC ACCTGTGCCT TCATATAGGG ATCTGGGTTC
1551 GTCTCTGCAG CTGCCTACCC CTGCCCCATG TGTCCCCCGT TACCTGTCCT
1601 CAGAGCCTCA GGTATGGTCC AGGCTCTGAG GAAAGTCAGG GTTGCTGTGT
1651 GGGAACCCCT CTGCCTGGCA CCTGGATACC CTGGGCCAGG TAACCTGAGG
1701 GCAGGGGAGA GGTGGGGTGG CAGACACGCA GAAGTGCTAC TAGTGACAGG
1751 GCTGCCATCG CTCACCTGTA CCTATTTACA CCCAGAACTT TCCAGCTCCC
1801 CCTCATCATG CCTCCTCCTT CCTACCTGCC TCCCCTCTGC TGGTGCACCT
1851 CGCCCAACTC ATTCTTACTG CACAGTTCAC TTTATTTAAC AATTTTCATG
1901 TCCCCCATCT CGCTCTGTGC CCCTCCCCAC CAGGGCTTCA GCAGGAGCCC
1951 TGGACTCATC ATCAATAAAC ACTGTTACAG CAAAAAAAAA AAAAAAAAAA
2001 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
2051 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA  (SEQ ID NO:1)
```

FEATURES:

5'UTR:	1-107
Start Codon:	108
Stop Codon:	1536
3'UTR:	1539

FIGURE 1A

HOMOLOGOUS PROTEINS:

Top BLAST Hits:

	Score	E
CRA|89000000199482 /altid=gi|7297404 /def=gb|AAF52663.1| (AE003...	330	4e-89
CRA|89000000199480 /altid=gi|7297402 /def=gb|AAF52661.1| (AE003...	330	4e-89
CRA|89000000199481 /altid=gi|7297403 /def=gb|AAF52662.1| (AE003...	330	4e-89
CRA|89000000197173 /altid=gi|7294781 /def=gb|AAF50116.1| (AE003...	268	1e-70
CRA|89000000195851 /altid=gi|7293314 /def=gb|AAF48694.1| (AE003...	265	1e-69
CRA|18000005127815 /altid=gi|7509795 /def=pir||T26845 hypotheti...	258	2e-67
CRA|89000000194855 /altid=gi|7292192 /def=gb|AAF47603.1| (AE003...	253	5e-66
CRA|89000000197171 /altid=gi|7294779 /def=gb|AAF50114.1| (AE003...	252	8e-66
CRA|89000000197172 /altid=gi|7294780 /def=gb|AAF50115.1| (AE003...	252	8e-66
CRA|18000005102492 /altid=gi|2429516 /def=gb|AAB71045.1| (AF025...	250	3e-65

BLAST dbEST hits:

	Score	E
gi|5422591 /dataset=dbest /taxon=9606 ...	1400	0.0
gi|3648072 /dataset=dbest /taxon=9606 ...	1017	0.0
gi|5746200 /dataset=dbest /taxon=9606 ...	730	0.0
gi|10249244 /dataset=dbest /taxon=96...	642	0.0
gi|8612353 /dataset=dbest /taxon=960...	329	1e-87
gi|10083945 /dataset=dbest /taxon=960...	313	7e-83

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|5422591 testis
gi|3648072 Testis
gi|5746200 Brain meningiomas
gi|10249244 Brain normal
gi|8612353 Head-neck
gi|10083945 Colon

From tissue screening panels:

Human whole liver

FIGURE 1B

```
  1 MSTQRLRNED YHDYSSTDVS PEESPSEGLN NLSSPGSYQR FGQSNSTTWF
 51 QTLIHLLKGN IGTGLLGLPL AVKNAGIVMG PISLLIIGIV AVHCMGILVK
101 CAHHFCRRLN KSFVDYGDTV MYGLESSPCS WLRNHAHWGR RVVDFFLIVT
151 QLGFCCVYFV FLADNFKQVI EAANGTTNNC HNNETVILTP TMDSRLYMLS
201 FLPFLVLLVF IRNLRALSIF SLLANITMLV SLVMIYQFIV QRIPDPSHLP
251 LVAPWKTYPL FFGTAIFSFE GIGMVLPLEN KMKDPRKFPL ILYLGMVIVT
301 ILYISLGCLG YLQFGANIQG SITLNLPNCW LYQSVKLLYS IGIFFTYALQ
351 FYVPAEIIIP FFVSRAPEHC ELVVDLFVRT VLVCLTCILA ILIPRLDLVI
401 SLVGSVSSSA LALIIPPLLE VTTFYSEGMS PLTIFKDALI SILGFVGFVV
451 GTYEALYELI QPSNAPIFIN STCAFI  (SEQ ID NO:2)
```

FEATURES:
Functional domains and key regions:
[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 7

1	31-34 NLSS
2	45-48 NSTT
3	110-113 NKSF
4	174-177 NGTT
5	183-186 NETV
6	225-228 NITM
7	470-473 NSTC

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2

1	3-5 TQR
2	334-336 SVK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 4

1	15-18 SSTD
2	20-23 SPEE
3	24-27 SPSE
4	112-115 SFVD

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

7-14 RNEDYHDY

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 7

1	42-47 GQSNST
2	59-64 GNIGTG
3	67-72 GLPLAV
4	175-180 GTTNNC
5	342-347 GIFFTY
6	404-409 GSVSSS
7	451-456 GTYEAL

[6] PDOC00009 PS00009 AMIDATION
Amidation site

138-141 WGRR

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	52	72	0.668	Putative
2	75	95	2.032	Certain
3	143	163	1.799	Certain

FIGURE 2A

4	193	213	1.467	Certain
5	216	236	1.884	Certain
6	258	278	1.566	Certain
7	289	309	2.126	Certain
8	335	355	1.378	Certain
9	375	395	1.332	Certain
10	398	418	1.748	Certain
11	437	457	1.533	Certain

FIGURE 2B

```
BLAST Alignment to Top Hit:
>CRA|89000000199482 /altid=gi|7297404 /def=gb|AAF52663.1| (AE003621)
          CG13384 gene product [alt 3] [Drosophila melanogaster]
          /org=Drosophila melanogaster /taxon=7227 /dataset=nraa
          /length=486
         Length = 486

 Score =  330 bits (837), Expect = 2e-89
 Identities = 184/425 (43%), Positives = 262/425 (61%), Gaps = 32/425 (7%)

Query: 47  TTWFQTLIHLLKGNIGTGLLGLPLAVKNAGIVMGPISLLIIGIVAVHCMGILVKCAHHFC 106
           T+ F TL+HLLKGNIGTG+L +P A KNAG+ +G     +I+G +  HCM +LV C+H  C
Sbjct: 78  TSNFDTLVHLLKGNIGTGILAMPDAFKNAGLYVGLFGTMIMGAICTHCMHMLVNCSHELC 137

Query: 107 RRLNKSFVDYGDTVMYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYFVFLADNF 166
           RR  +  +D+ +       ES P   LR ++   RR+V  FL +TQ+GFCCVYF+F+A N
Sbjct: 138 RRFQQPSLDFSEVAYCSFESGPLG-LRRYSMLARRIVTTFLFITQIGFCCVYFLFVALNI 196

Query: 167 KQVIEAANGTTNNCHNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANI 226
           K V++          H  +       M  ++Y+L  L  ++LL  +RNL+ L+  SL+A +
Sbjct: 197 KDVMD---------HYYK-------MPVQIYLLIMLGPMILLNLVRNLKYLTPVSLVAAL 240

Query: 227 TMLVSLVMIYQFIVQRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPR 286
             +  L + + +++  +PD   +  VA W T PL+FGTAI++FEGIG+VLPLEN M+ P
Sbjct: 241 LTVAGLAITFSYMLVDLPDVHTVKPVATWATLPLYFGTAIYAFEGIGVVLPLENNMRTPE 300

Query: 287 KF---PLILYLGMVIVTILYISLGCLGYLQFGANIQGSITLNLP-NCWLYQSVKLLYSIG 342
            F     +L  GMVIV  LY ++G  GYL++G +++GSITLNLP    L Q V++  ++
Sbjct: 301 DFGGTTGVLNTGMVIVACLYTAVGFFGYLKYGEHVEGSITLNLPQGDTLSQLVRISMAVA 360

Query: 343 IFFTYALQFYVPAEIIIPFF-----VSRAPEHCELVVDLFVRTVLVCLTCILAILIPRLD 397
           IF +Y LQFYVP  I+ PF       +RA +    V    +R VLV  T +LA  IP L
Sbjct: 361 IFLSYTLQFYVPVNIVEPFVRSHFDTTRAKDLSATV----LRVVLVTFTFLLATCIPNLG 416

Query: 398 LVISLVGSVSSSALALIIPPLLEVTTFYSEGMSPLT--IFKDALISILGFVGFVVGTYEA 455
            +ISLVG+VSSSALALI PP++EV TFY+ G        ++KD LI I G  GFV GT+ +
Sbjct: 417 SIISLVGAVSSSALALITAPPIIEVITFYNVGYGRFNWMLWKDVLILIFGLCGFVFGTWAS 476

Query: 456 LYELI 460
           L +++
Sbjct: 477 LAQIL 481  (SEQ ID NO:4)


>CRA|89000000199481 /altid=gi|7297403 /def=gb|AAF52662.1| (AE003621)
          CG13384 gene product [alt 2] [Drosophila melanogaster]
          /org=Drosophila melanogaster /taxon=7227 /dataset=nraa
          /length=482
         Length = 482

 Score =  330 bits (837), Expect = 2e-89
 Identities = 184/425 (43%), Positives = 262/425 (61%), Gaps = 32/425 (7%)

Query: 47  TTWFQTLIHLLKGNIGTGLLGLPLAVKNAGIVMGPISLLIIGIVAVHCMGILVKCAHHFC 106
           T+ F TL+HLLKGNIGTG+L +P A KNAG+ +G     +I+G +  HCM +LV C+H  C
Sbjct: 74  TSNFDTLVHLLKGNIGTGILAMPDAFKNAGLYVGLFGTMIMGAICTHCMHMLVNCSHELC 133

Query: 107 RRLNKSFVDYGDTVMYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYFVFLADNF 166
           RR  +  +D+ +       ES P   LR ++   RR+V  FL +TQ+GFCCVYF+F+A N
Sbjct: 134 RRFQQPSLDFSEVAYCSFESGPLG-LRRYSMLARRIVTTFLFITQIGFCCVYFLFVALNI 192

Query: 167 KQVIEAANGTTNNCHNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANI 226
           K V++          H  +       M  ++Y+L  L  ++LL  +RNL+ L+  SL+A +
Sbjct: 193 KDVMD---------HYYK-------MPVQIYLLIMLGPMILLNLVRNLKYLTPVSLVAAL 236

Query: 227 TMLVSLVMIYQFIVQRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPR 286
             +  L + + +++  +PD   +  VA W T PL+FGTAI++FEGIG+VLPLEN M+ P
Sbjct: 237 LTVAGLAITFSYMLVDLPDVHTVKPVATWATLPLYFGTAIYAFEGIGVVLPLENNMRTPE 296
```

FIGURE 2C

```
Query: 287 KF---PLILYLGMVIVTILYISLGCLGYLQFGANIQGSITLNLP-NCWLYQSVKLLYSIG 342
           F     +L  GMVIV  LY ++G  GYL++G +++GSITLNLP    L Q V++  ++
Sbjct: 297 DFGGTTGVLNTGMVIVACLYTAVGFFGYLKYGEHVEGSITLNLPQGDTLSQLVRISMAVA 356

Query: 343 IFFTYALQFYVPAEIIIPFF-----VSRAPEHCELVVDLFVRTVLVCLTCILAILIPRLD 397
           IF +Y LQFYVP  I+ PF       +RA +    V    +R VLV  T +LA  IP L
Sbjct: 357 IFLSYTLQFYVPVNIVEPFVRSHFDTTRAKDLSATV----LRVVLVTFTFLLATCIPNLG 412

Query: 398 LVISLVGSVSSSALALIIPPLLEVTTFYSEGMSPLT--IFKDALISILGFVGFVVGTYEA 455
            +ISLVG+VSSSALALI PP++EV TFY+ G       ++KD LI I G  GFV GT+ +
Sbjct: 413 SIISLVGAVSSSALALIAPPIIEVITFYNVGYGRFNWMLWKDVLILIFGLCGFVFGTWAS 472

Query: 456 LYELI 460
           L +++
Sbjct: 473 LAQIL 477  (SEQ ID NO:5)


>CRA|18000005127815 /altid=gi|7509795 /def=pir||T26845 hypothetical
          protein Y43F4B.7 - Caenorhabditis elegans
          /org=Caenorhabditis elegans /taxon=6239 /dataset=nraa
          /length=607
         Length = 607

 Score =  258 bits (652), Expect = 1e-67
 Identities = 142/418 (33%), Positives = 235/418 (55%), Gaps = 19/418 (4%)

Query: 40  RFGQSNSTTWFQTLIHLLKGNIGTGLLGLPLAVKNAGIVMGPISLLIIGIVAVHCMGILV 99
           R    NS T  Q  IH++K  +GTGLL LPLA K++G+ +G I  ++I ++ ++CM  +V
Sbjct: 42  RLPTENSLTPEQAFIHMVKAMLGTGLLSLPLAFKHSGLFLGLILTVLICLICLYCMRQVV 101

Query: 100 KCAHHFCRRLNKSFVDYGDTVMYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYF 159
             AH  C R  +  +DY + +   +E  P  W++ + ++ +++V+  + ++QLGFCCVYF
Sbjct: 102 FAAHFVCNRNGRDLIDYANIMRGAVEMGP-PWIKRNGYFFKQLVNVNMFISQLGFCCVYF 160

Query: 160 VFLADNFKQVIEAANGTTNNCHNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSI 219
           VF+ADN +                NN T I    +   ++ML  L  ++ +  IR L  L+
Sbjct: 161 VFMADNLEDFF----------NNNTSI---HLSKAVWMLLLLIPMLSICSIRRLSILAP 206

Query: 220 FSLLANITMLVSLVMIYQFIVQRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLE 279
           F++ AN+  +V++ ++  F +  +   S LP         PLFFGT +F+FEG+ +++P+E
Sbjct: 207 FAMAANVVYVVAVAVVLFFFLSDLRPISSLPWFGKATDLPLFFGTVMFAFEGVAVIMPIE 266

Query: 280 NKMKDPRKFPL---ILYLGMVIVTILYISLGCLGYLQFGANIQGSITLNLPNCWLYQSVK 336
           N+M+ P  F     +L    ++V  ++   G  GYL  G +++ + TLNLP     YQ++K
Sbjct: 267 NRMQSPHAFISWNGVLNSSCLVVLAIFSVTGFYGYLSLGNDVKDTATLNLPMTPFYQTIK 326

Query: 337 LLYSIGIFFTYALQFYVPAEIIIPFFVSRAPEHCELVVDLFVRTVLVCLTCILAILIPRL 396
           LL++  I  +Y LQFYVP E I  +   + P   + +     R   V LTC +A LIP L
Sbjct: 327 LMFVACIMISYPLQFYVPMERIEKWITRKIPVDKQTLYIYIARYSGVILTCAIAELIPHL 386

Query: 397 DLVISLVGSVSSSALALIIPPLLEVTTFYSEG-MSPLTIFKDALISILGFVGFVVGTY 453
            L ISL+G+ S +++AL+ PP +E+ T Y++  +S      K+ ++    F+GF  GTY
Sbjct: 387 ALFISLIGAFSGASMALLFPPCIELLTSYAKNELSTGLWIKNIVLLTFAFIGFTTGTY 444  (SEQ
ID NO:6)


>CRA|335001101719045 /dataset=FastAlert /length=476
          /altid=Derwent|WO200071709.21
         Length = 476
 Score =  909 bits (2324), Expect = 0.0
 Identities = 450/476 (94%), Positives = 459/476 (95%)
 Frame = +3

Query: 108  MSTQRLRNEDYHDYSSTDVSPEESPSEGLNNLSSPGSYQRFGQSNSTTWFQTLIHLLKGN 287
           MSTQRLRNEDYHDYSSTDVSPEESPSEGLNNLSSPGSYQRFGQSNSTTWFQTLIHLLKGN
Sbjct: 1    MSTQRLRNEDYHDYSSTDVSPEESPSEGLNNLSSPGSYQRFGQSNSTTWFQTLIHLLKGN 60
```

FIGURE 2D

```
Query: 288  IGTGLLGLPLAVKNAGIVMGPISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFVDYGDTV 467
            IGTGLLGLPLAVKNAGIVMGPISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFVDYGDTV
Sbjct: 61   IGTGLLGLPLAVKNAGIVMGPISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFVDYGDTV 120

Query: 468  MYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYFVFLADNFKQVIEAANGTTNNC 647
            MYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYFVFLADNFKQVIEAANGTTNNC
Sbjct: 121  MYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYFVFLADNFKQVIEAANGTTNNC 180

Query: 648  HNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANITMLVSLVMIYQFIV 827
            HNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANITMLVSLVMIYQFIV
Sbjct: 181  HNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANITMLVSLVMIYQFIV 240

Query: 828  QRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPRKFPLILYLGMVIVT 1007
            QRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPRKFPLILYLGMVIVT
Sbjct: 241  QRIPDPSHLPLVAPWKTYPLFFGTAIFSFEGIGMVLPLENKMKDPRKFPLILYLGMVIVT 300

Query: 1008 ILYISLGCLGYLQFGANIQGSITLNLPNCWLYQSVKLLYSIGIFFTYALQFYVPAEIIIP 1187
            ILYISLGCLGYLQFGANIQGSITLNLPNCWLYQSV+LLY  GI  TY LQFYV A+II+P
Sbjct: 301  ILYISLGCLGYLQFGANIQGSITLNLPNCWLYQSVELLYLGGICLTYPLQFYVSAKIIVP 360

Query: 1188 FFVSRAPEHCELVVDLFVRTVLVCLTCILAILIPRLDLVISLVGSVSSSALALIIPPLLE 1367
              VS   + C L+VDL + + ++C TCILAILIPRLDLVISLVGSVSSSALALIIPPLLE
Sbjct: 361  VIVSWVCKCCTLMVDLGIGSAMLCKTCILAILIPRLDLVISLVGSVSSSALALIIPPLLE 420

Query: 1368 VTTFYSEGMSPLTIFKDALISILGFVGFVVGTYEALYELIQPSNAPIFINSTCAFI 1535
            VTTFYSEGMSPLTIFKDALISILGFVGFVVGTYEALYELIQPSNAPIFINSTCAFI
Sbjct: 421  VTTFYSEGMSPLTIFKDALISILGFVGFVVGTYEALYELIQPSNAPIFINSTCAFI 476  (SEQ
ID NO:7)
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF01490	Transmembrane amino acid transporter protein	223.3	3.5e-63	1
PF01091	PTN/MK heparin-binding protein family	2.0	9.5	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF01091	1/1	192	208	..	1	17	[.	2.0	9.5
PF01490	1/1	71	451	..	1	467	[]	223.3	3.5e-63

FIGURE 2E

```
   1 AAAACCAGAA AGTCAGATAG TCCCTGTCTC ATCTTCAATC TCTTTATTTG
  51 TTATTAGTCT GTTCAGGCTT TCCATTTCTT CCTGATTTCA ATCTTGGTAG
 101 GTTGTATTTT TCTAGGGATT TCTCCATTTC ATCTGGGTTA TCCAATATGT
 151 TGGCAAATAA TTGTTCACAA TAGCCCCATA TGATTCTTTT TATTTCTGAA
 201 GCTTCTGTTG TAGTGTCTTC ACTTTCATTT TTTATTTTAT TAGTCTTCTT
 251 TTTTTTCTTA GACTAGCAAA GGGTACGTCA ATTTAATTTT TTCCAAAAAA
 301 TCAACTCTAG TTTTATTGAT TTTGTCTGTT TTTTTTTCTG TTGTCTATTT
 351 GATTTATTTC TGTTCTGCTC TTTCTTTTCT TCTTTTAACT TTGGGTCTTG
 401 TGAGCTCTTC TTTTTCCAGT TTCCTGAGGT ATAATGTTAA ACTATTTAAT
 451 AGGTCTCTTT CTTCTCTTTT AATGTAGGCA TTTATGTCTA TAAACTTCTC
 501 TCTTAGAACT ACTTTTGCTG CATTCCATAA GCTTTGGTAT GTTATGTCTC
 551 CATAGGCAAT TGTCTGAAAA TATTTTTTAA ATTTCCCGTT TGATTTCATC
 601 TTTGACCCAC TGGTTGTTTC GAAGCATAGG ACTGGGTATG GTGGCTCACA
 651 CCTATAATCT CAGCACTTTG GGAGGCCGAG GTGGGCAGAT CACCTGAGGT
 701 CAGGAGTTTG AGACCGCCTG ACCAACATGG TGAACCTCGT CTCTACTAAA
 751 AATACAAAAT TAGTCGCGCG TGGTGGCACG TGCCTGTAAT CCCAGCTACT
 801 TGGGAGGCTG AGACAGGAGA ATCGCTTGAA TCCAGGAGGC AGAGGTTGCA
 851 GTGAGCCAAG ATTGCACCAC TGCACTCCAG ATGGGGCAAT AAGAGCGCAA
 901 CTTTGTCTCA AAGAAAAAAA AAAGCGTGTT GTTTAATTTC CACATATTTG
 951 TGAATTTTCC AAGATTCCTC CTGTCATTGA TTTCTAGCTT CATATTATTA
1001 TGGTCTGAAA GAATATTAAT ATGATTTCAA TCTTCTTAAA TTTAAGGCTT
1051 GTTTTTTGGA CTAGCATATG GTCTATTCTA GAGAATGTTT CAAGTGTGTT
1101 AGAGAAAAAA TGTGTATTCT GTTTCTGTTG AATGGAAAGT TCTGTATATA
1151 TCTGTTAGGT TCATTTGGTT TAAAGTGCAA TTCAAGTTCA TTATTTCCTT
1201 ATTTTCTCTC TAGTTGCTCT ATCCATTGTT GAAAGTGGGA TATTGACCCT
1251 CCTACTATTG TGTTGCTATC TATTTCTCCC TTCATGGCCA TTAATATTAG
1301 TTGTATGTAT TTAGGTGCTC CAATTTTGCA TGCATATATA TTTACAGTTG
1351 TGTCTTCTTG ATGAATTGAC CCCTTTATTA TTAAACAATG ATCTTCTCTG
1401 TCTCTTGTGA CAGTTTTTGA CTTGAAGCCT ATTTGTTACA TAATTGTTAA
1451 AGGAAAAGTC TGTAACAAGG AGGTAAAAGG AGAAGCCTAG ATAATACAAT
1501 ACTGAAATGT TGCCATCCAT TTAAAAATGT TACTTTAAAA ATTTGAATGT
1551 ATTAAAAGAT AATGTTGCCC TACCCCACAG TTCCATTTCC AGTAGCAACC
1601 ACAGATGATA GTTTGTGTAT GCTTCTGAAA AATTGGAAGT TTTAAAAATA
1651 TGCATATTTC TTATTATAAA AGCAATACAA ACTCATCGAG ATGTGTAAAA
1701 GAAAATACAG CCAGTGTAAA AATTAGCAAT ATTTCACAAA CCCACAACTC
1751 AAGGGACAGT GCTCTTCGAC TGGACTCTGC CCCATGCCCA AGATCAATGC
1801 CCTGTTCAGT TCCTATTCGC AGTCCCCAGC GCCCAGGAAC ATAGTCCTTC
1851 CAGCAGTGGC AGTAATAGGT CGCCAGGTGG TGCTGTGGAG CAGAGCTCCG
1901 GAGCTCAGTG AGAAAAAAGG CGCGGCCGCT CAAGGGAGCA CGTGACCTCG
1951 GCCTCTGGCG TGGGCGGTGG GATCACGTGA TGAGGTCCGG AAGCGGCTGC
2001 CGGGCAGCAA AGGAGGATGG CGAGGGGCTG ATACTGAACC CGGGAAGGGT
2051 GGGCTGTGCT GAAGCCAGAG CCGGAGCCGG AGCTGGGGCC AGAACCCGAG
2101 CAGTGAGTTC CTCCACTGAC GAGTTCCGGC TGGCGGCGCT CGCCGCCTTG
2151 GGCAGGACCC ACCTCGCCTT CCTCCCGGCG TGGCAGATGC TCCAGGTCAG
2201 GCACTGGATC CGCCCGGGCT GTGGGTCCGC GACTCCTTGG CGTCCCCGGG
2251 CCGCAGCTGC GGTACGACGC TGACACCCCT CTGTGAATTG GGCGAAGCGT
2301 GGAGATCCCT TGTCCCTCGC GCTATCTCCC TTGACCTCGT GGGGTTGGGA
2351 TCTCACCGTC CTGTTTGACT GACAGGTGGG GGAAACTGGG GTAGATGGTG
2401 AAGATAACCC AAAGGACCAT CTAGGGCGTC TTTCACGCTT CGCACAGGTC
2451 TCCCCGTTTC CAGCAAATGT CTTGCCCGCT GCGGGAGCGC TGCTTGAGAC
2501 AGGCTCATAA TGGGTCTTTG GGTCAGAACT GCAAGGACGC TGGGAAGTCG
2551 TCTGGTGCAG CTCCCTCCTA GGACAGTTGG AGAAACTGAG CCCTTACTCC
2601 GGGAAGGGGT AAGGGCTTGC CTAAGGTCAT CCAGTGAGTT AATCGGAGAC
2651 CCGGAGACCT GCGACTAGAA TGCAAATGTT CCTAAGCTTC AGCAGCTGTT
2701 TGCTTTTCGC CACACCGCCT CCTGCGGGAA ACTTCACCTG TGAAAAGGCA
2751 CTCCTTTCTG TCCCTTTCTC TTTTAGTCCT CTCCCTTTTT AGCTGTCTGC
2801 ATTTTCCACC GCTGGGGTTG GATTTGCTCT GGGTGTGGTT CCCTGTTTGT
2851 TCATTATTTT TCTGCAAACT CATCCTTCTG TAGGTTTGGT TTCTAACCTT
2901 CCTGCATTCT ATGTAAGTCA CACCAAAATA TGAAATATGA ATCGGAATGT
2951 GCTTCTGGGA AGATAGGTGG CTGAGCCGAG GTTGTGGAGA GCCCTGACGT
3001 TAACTTGAAG AATGTAAAGA CCTTTGCTTT ATTTTTTCTG TAACTTGTCA
3051 GATTTGGGAT TGCTTATTTG GATGGACGTT TTGCAGTTAT TTGAATTTTG
3101 CTGAAGATAG CATCATGGTG CAATGGACAG AACAGAGATT GGGGAATCAG
3151 GATATTTTGT CCTAGCTCTG CCGCTTACCT GGCAACCTTA AGTGACTCGC
3201 GTTTGGGTTT CTCAGTCTAG ACAGTGATGG AATTGAATTC TTAAGGGCCC
3251 CTTCTGCTGT GATCTGGATG TTGTGCATCT TTCTAGGCTT GTTTTTTTGT
3301 TTGTTTGTTT TTAAATAGAG ATGAGGTCTC ACTATGCTGC CCAGGCTGAT
3351 CTCAAACTCC TGGGCTCAAG TGATCCTCCC ACCTTGGCCT CCCAAAGTAT
```

FIGURE 3A

3401 TGCGATTACA GGGGTGTGAG CCAGTGCCCC TGACCAGGGT CTGTTTGTTT
3451 TTTTATTCCG AGAGATTTTA CCCGCTGTGT ACACTGAGTA TCAGCCTTGC
3501 AACAAGACTT AATCTAATTG TGTAGGAAGC AGTTTCCTCT GCTTATTCCT
3551 CTGTTGCTAT AAAATCCTCC TCCTCTTTCT TCCTATCTCT GTATTATGTC
3601 TAAGCTAAAT ACTAACAGCT GAAAATGATT TTTAACTGTC TGTTATTATT
3651 TTAAACATGA TCAGGGCCCT CTTCTGACTC TTGTCTAGAG CCTCGTTTAC
3701 AGTAATTAAC TTACTTGTAC ATTTAGAACA TGTTTTCGTT AAAAATGTTC
3751 TTGAAGTCAA GTGGAAAGGG ACAAACTTTG TGTTTGTCTG GGGATGGAGG
3801 CTACGTGCAG CAGAGGCTGT GCTGATAACA GCTTTAAGCC TCTCCCTTTG
3851 TTTTCTGATT GTATCGTTTA TATTCGCCTG CTCTGGAGTC CTTGTTTTCC
3901 CTTTAAGGCA TGTAGCTGTT CATTCAGCCT CCTCTTGTGG CAGTTCGAAG
3951 TGCTCAGAGT TTTCTCTGTT CCAAAGGCGG TGTAAGAAAA AGGATTCCAT
4001 TTATTTTATA ATATTCTGGA GACTACAACT GGGACCAATA GATGAAAGTG
4051 CCATAAGGGA AGGGTGTTTG GTAAGTGTTA AAAACTTTCA AAGAATTAGT
4101 GGTCTCCTGT ATTAGAACAC TGAGCACGGA GCTCCCTGTC ACTGGAGGTA
4151 ATCGTCTCTG GGCAATTACT TCGGGGATGT TGTAGATGAA ATTAGATTGT
4201 TGGGTAGAGG GTTGGACTAG ACAAAATTTT AAGTTTTCTT TCAACTCAAG
4251 AGTCTGTGAC ATTCTAGGAC TGGACTTACT AGCATGTAGA GTGGATGGAG
4301 CAGATGTCCA CTTACTAGCA TGTGGGATGG ATGGAGCAGA TGTCCACTTA
4351 TTACTAGCAT GTGGGATGGA TGGAGCAGGT GTCCACTTAG TGGCATGTAG
4401 AGTGGATGGA GCAGATGTCT GCTTACTAGC ATGTGGGATG AATGGAGCCG
4451 ATGTCCACTT ACTGGCATGT AGAGTGGATG GAGCAGATGT CCACTTATTA
4501 GCATGTGGGA TGGATGGAGC CGATGTCCAC TTACTAGCAT GTAAAGTGGA
4551 TGGAGCAGAT GTCCGCTTAC TAGCACGTAG AGTGGACGGA GCAGATGTCC
4601 ACTTACTAGC ATGTGGGATG GATGGAGCTG ATGTCCACTT ACTAGCATGT
4651 GGGATGGATG GAGCTGATGT CCACTTACTA GCATGTGGGA TGGACGGAGC
4701 CGGTGTCCAC TTACTAGCAT GTGGGATGGA CGGAGCCGGT GTCCACTTAC
4751 TAGCATGTGG GATGGATGGG GCAGGTGTCC ACTTACTAGC ATGTAGAGTG
4801 GATGGAGCAG ATGTCCACTT ACTAGCATGT AGAGTGGATG GAGCAGATGT
4851 CCACTTACTA GCATGTGGGA TGGACGGAAC TGGTGTCCAC TTAACTAGCA
4901 TGTGGGATGG ATGGAGCCGA TGTCCACTTA CTAGCATGTG GGATGGATGG
4951 AGCAGATGTC CACTTACTGG CATGTAGAGT GGATGGAGCT GATGTCCAGT
5001 TTTGTGATTA CTTTGTTTCT ATTTATAACC TTGTCTCAGG TAACTATTCT
5051 CATATTAAGT ACTCCTGCTT TTTCTTTCTT TTTATCACCA CCACCTCCCC
5101 TCCAGTGAGT ATCTCAGTTC TTTAAATGCT TGATATCGCT TCAAAGGTCA
5151 GATGAGTGAA TAGTCTTCTG TTTCTGCTTT TCCTGGCCTG GTGCTGATAA
5201 CCGCTTCCAA AGTGCATGAC TGATTAGCAT TACTCACACC TAGGCCAGCT
5251 TTTCTCTTTT TCCCATAGAG GAACTCACAT GGAATCCGTT TATTTCCATC
5301 CAGGCCTTCT CTTGTTTCCC ATAGAAGAAG CTCTCACTGA GTCTGTTTAC
5351 TTCCACCCAG GCCTTGGAAG AATCCTGTAC CTCTCTCCTT TGGCCAGGCC
5401 TTACTGTGAT GAGCACATAA AGGTAGCCTC TACTTAATGG GCATGGGGGC
5451 CGATGGATGG GGCATTGTAA ATAGGCTGAA ATAGGAACCA CACGGTGCTG
5501 CATTTGGGGT TGTCTCTTCT TTTATCCCCC AAAATATTTC TCTTGGAAGC
5551 CTTGACACCC AGGGCCAGTT CTTTTTTTAC TTATTATTTA TTTTATTTTA
5601 TTATTATTAT TTTTTTGGAG ATGAGGTCTC ACTATTTTGT CCACGCTGGT
5651 CTTGAACTCC TGAGCTCAAG CAGTCCTCCC ACCTTGGCCT CCCAAAGTGC
5701 CGGGATTACA GATGTGAGCC ACCATGCTTG GCCCCAGGGC TAGTCCTTAA
5751 TGACCATCTC TAGTAGGAAA AGCCACTCTG TGCTTTCCTT TTCCATGAAG
5801 TTAGGAAATC TGTCTGTGGG TTACTGAGAT GTTCATGTCA CTTAGATCAC
5851 CATCTCCAAG GTAGGGACCT GGCTTCACAA TCCAGAGTTT TAAATGGAGC
5901 CCATACTGCA GACTTTGTTT AGTGAACTTT CTCACTTTCT GTCCTTGAAC
5951 TTCTCTGTAG TAATAATCAC AATTGCTATC ATTAATGAGG GTTTATTATG
6001 TTCCAGGCAA CATATCTAAC ATTTATTTAT TTTTTCCTCC TATCTTCATA
6051 ACAATATTGT GATGTAGATG TTATTAATGA CATCTTTCAG ATGAGGAGAC
6101 TGTGGCAAAG GGAGATGAAT TAACTTGCTC AGAGTCACAC TACCAGACTG
6151 CAAACTCAGG TGCTTTTTAT TGTGCAGAAT ACCCGCTGCA GACCTAATCC
6201 TGCCCCAGGC TCTGGGGCCA GCTTTGTTCG CAGGGAGATT TTAAGGAGGG
6251 TATATAATTT AAGGTGTGGT AGAAAGAATA CTGGACTGGG ATGCTGGTTT
6301 GCTGGACTGT CATCTCAAAT CTTTGATTTG ACTCATCCTG GGGCCTGGAT
6351 GAGTCAGCCT TTGTGTGTGG GCCCTGGTTT TCTGACCCCT AAGAAGGAAG
6401 CTGGAGCTTG ACCTTCTCTA AAGCTATACC TGGCCCTAAC ATTTAGTGAT
6451 CTTCATGGTT GGGAGTAAAA GTGTGCGTGT TTGCCTGTTC AGCAGCTGCT
6501 TTGTGCAGAA CCTGCTGAGG TCAGCAGCTG CCCTGTAGCT GTTCTAGCAT
6551 CAGACTCCTA CAGGAAAAAG TCTCAATTTA TGGAATGTTC TGCTCTGGTA
6601 AGTTGGATGG AATTCTATCT GATGCTGTTT TAAAAACAAA TTATGTAGAA
6651 GCCAAACCAT TTTACTTCCC TCACTGTAGA CCACACATAG CAACACAGTC
6701 TGTGTCTTTG TTCATGTTTT TAGAATTCCA TCGACAGAGA GGAGAAAATA
6751 CATCTGGGGA ATTTGCCGCT GCTCTGAGTT CCAAAGTCCA AACCAATGTA

FIGURE 3B

6801 ATTGTTTCAG AATAACGGAT GACACTTTTA GCTTGCAAAC AAGGGGCGCC
6851 AATGCGTGAA TTCTGGTAGG AGGTGAGGCC TAGGGTGTAC CTATCATAAT
6901 AAGATCATAT ATTTTTTGTA GTGCTTTATA TAAATCTACC TATAATCAAG
6951 ATTACCTAGG AAGCTAGTTA AAAATAAAAC GCCTCTTGCC TGTAATCCCA
7001 TCACTTTGAG AGGCTGAGAC AGGTGGATCC CTTGAGGTCA AGAGTTTGAG
7051 ACCAGCCTGG CCAACACGGA GAAACTCCAT CTCTACTAAA AACACAAAAA
7101 ATTATCTGGG CATGGTGATG GACGCCTGTA ATCCCAGCCA CTCGGGAGGC
7151 TGAGGCAGGA GAATCGCTTG AACCCAGGAG GCGGAGGTTG CAGTGAGCCA
7201 AGATCACACC ATTGCACTCC AGCCTGGGCA ACAGAGGGAG ACACCATCTC
7251 AAAAAAAAAA AAAGAAGACA AAAAGACAAA AACAACAACA AAAAAACATA
7301 GGCTGGGCAT GGTGACTCAT GCCTGTAATC CCAGCACTTT GGGAGGCCAA
7351 GGTGGGTGGA CCACCTGAGG TCAAGAGTTT GACACCAGCC TGGCCAACAT
7401 GATGAAACCC CGTCTCTACT AAAAATAGAA AAAAATTAGC CAGTTGTGGT
7451 GGCGCATGTC CGTAATCCCA GCTACTCGGG AGGCTAAGAC AGGAGAATTG
7501 CTTGAACCTG GGAGGCGGAG GTTGCAGCGA GCCAAGATCG CACCACTGCA
7551 CTCCAGCCTG GGCAACAAGA ATGAAACTCC ATCTCCAGTA AATAAATTAA
7601 AATAAATAAA TAAAATAAAA TAAAATAAAA TGCTAAGGTG GAATCAAGTT
7651 GGGCCCAGAA ATCTATTTTT TTTTTCCTTG ACGTATGTTT CATTTAACCC
7701 AATATATCCC AGATATTATC ATTGCAATAT ATAATCAGTA TAAAGATTAT
7751 TAATTCATGG GATATTTCAC AATTTTTTTG TTACCAGTTC ATTGAAATCT
7801 AGTGTGCACA TTTCAATTTT ACCCAAGTGT ATTTCAAGTG TAAGATAGCT
7851 ATTTATGGCT AGTGGTTACT GTACTGGATG GTACAACTTC AGAATATGTT
7901 ACCATCTATT GATCTTAATC CTCCTTTATT TTGAACAAAC CCAGTCACTA
7951 AAAAATTGAA ATTGGAATCC TGAAACTTTA GAAGTGAAAG TGTACTTAGA
8001 AATCATCTAA TGCAGTTTTC TCAATTCTAT ATCAAAATAA GAAAACAACT
8051 TTGGGATTAG AATGACAGCC AGATTATGTT CTCCTGAGTC CTGAATCCCA
8101 TGCTGTTAAA ATGGGAACAT TAGCATTTGA ATTTATTAGA AAAATTTCTG
8151 GCCTTGCCTT AAAAAAAAAA AATCACTGTA GAATTCCCCT TAAAATTGCC
8201 CACTTCTGAA AAATTTAACA CCTACAAATT TTTATTTTTA AAAATAGAAT
8251 AAAATTTATT TTATTTTTAA AAATAAAAAT TCAGTTTGCA CATACATTTT
8301 CCATATTGCA TCCGTTGCAC AAAGTGATTC CACCTGCTCA TTTTTAGTGC
8351 CCATCTAAAA ATGGCATATT TTGTAGATTG AAGAGCAACA CTTGTCTATT
8401 TATACAGCTA AAACAATAGT TACATAAGGA AAAAAAAGGA ATGTTTTAAG
8451 GTTTGTACAC TTAAATTTTT TTTTTTTTTT TTTTTTTTTG GCCATCAAAC
8501 TTGCAGACTT TTTTTACTCA GTTGCTCACT CTTCTGAGTC TAAATATCTA
8551 ATGGAGATTT GGACTTTGTG TTCTGTTTAT TGTCCTCAGT AATCTGAAGG
8601 ACAAGCTTGC CTTCAACTCT CACATAGTAC AACCCTCATT TAGACAGTTA
8651 ACAGGTACTA TTAAAATCTC CCATAGGGCG GGAACTGGCA ATTGCAGCAA
8701 TAGACTTGGC TATCAGATTT CATCAAAGGG AGCCTAAGGG CAGTGTGGCC
8751 ATGGATGCCA GCACTCATGG GGACAGACAG AGAGCAGGAG GAGGAGGCCT
8801 TGGTTTCCAA AAAGAGCCAT AGAAAGAACT CCGGGGAGTG GCTCTGCCCA
8851 CTGTCTGATG CTTGAATCCT TACATAACTG CTCTGAGAAA GGGCTTTTGC
8901 TTGGATTTTT TCAGGGATAA GGGAACAGGC TTTCTCCCAG AGTGATCTGT
8951 TCTATTTGGA ACAGATCTGT CTTTGATAGA AAGTTCTTCC TTACACCTAG
9001 CAAAAAATCA GCCCTCTTGA CTCTCCACGT ACTGATCCTA GCCCTGCCTG
9051 ACCTTTGAGG CCCCAAATAA CAAGTCTAAT CCATGTGACA GCTTTTTTTT
9101 TTTTTTTTTT TTGAGACGGA GTCTCGCTCT GTCGCCCAGG CTGGAGTGCA
9151 GTGGCGCGAT CTCGGCTCAC TGCAAGCTCC GCCTCCCGGG TTCACACCAT
9201 TCTCCTGCCT CAGCCTCCCG AGTAGCTGGG ACTACAGGCG CCCGCTACCA
9251 CGCCCGGCTA ATTTTTTGTA TTTTTAGTAG AGACGGCTTT TAAAGACAGT
9301 TTTTGTACCC CTCAAGTTGC TAGGTGGAAC CTTCTCAGTG CTTTCAACCA
9351 TTCCTCATTT AGTTGGTTTC CTACCCCTCT TGATCCTAGT TCTGACCCCT
9401 GGATATACCA CAATTTGTCA TTATCCCCTT TATAGCATGC TGCCTGGAAG
9451 AGAACACATT ATCTGGCAAT TCTGAGTTGT GTAACATGTA CCCATGTGTA
9501 ACATCTAGCA TGTAGAGTGG ATGGAGCAGA TGTCCACTTA CTGGCCTGTA
9551 GAGTGGATGG AGCCGAGGTC CACTTACTAG CATGTGGGAT GGATGGAGCC
9601 GATGTCTACC CATGATTCTC TCATGTCCTA ATGCAACCTA GAATTGTGTT
9651 GGTTTATTTG GCATCTTGGA TTATATTATT GCCTTCTGTT GAGCTTATCA
9701 TCAACCAGAA CTCCCAAGCA GAATTTTTTT TTTCTGTTTT CAATATGCAT
9751 GCAGTTGCTT AGCCACATCT TCCATATCCC GCAGGCTTAT TGGACCTTAA
9801 ATCTATAGAT TTCAGCTTCT TGTTGAGAAG TTTTCTTAGC AATATTCTGC
9851 TGTGGGTAAC TTCAGTTTTT ATACACAACA TAAAGAAGTC TCTCTGCAGT
9901 GTTTGAGATA AATTGAACAT CTGTACCAAG TAGACAACAG AGAGGTTTCT
9951 CGGTTGCTAG GGAAGGATTG GGCAATTAAT AAGTCCCTGT ATTCCATCCT
10001 TTCACCTTCA GTAATATATA GGTGTCAACC TAAAGGAAGA AGTTGAGACA
10051 CAAAATGCAA TTTTTAACAG TTTACTTGAA CTGTTTACTT GAACCAAGTG
10101 AGGACAGCTG CCCGGGACAC ACTTCCAAGT TGCCTGGGGG AGTGCGTCCT
10151 TCGGCCTTTG TTACCACAGA TTCTTAAAGG CAAAAGCGAA CAAGGAGAGG

FIGURE 3C

10201 ACTGATACAA AGTGACTTGA CAGGAATTCT CATCAGTTTA CAGAAATAGC
10251 ATGGATTATT GATGGGCTGT ACATTGTTGG ACTATAGGGT ATGAGTTATG
10301 ATGTCCAGTG TTAGCATTTT ATGACTTAGT GGTGTCAGTT AGTCTAGAAC
10351 CCACATAGCA AGTGGCTTCA AGAGGTAATT ATTTAACTCA AGGGGGGAGT
10401 GACACATGAC TGCTCTCACA TTTTAGTGCC TCTCTGGACC CGTAATTTAA
10451 AGGGATTCCT CAGATAAAAA GTTTCTTTTC TTTCTCACAA GATTCACTTG
10501 GAAGGTTCTA TCTTCAGATG CTTTTGGTTT GTTGGAAGGG ACTAGAAATT
10551 GGCAGCTTTT TCTTTTTTTC AGTAGAGGCA GGGTCTCACT ATGGTGCCCA
10601 GGCTGGTCTT GAACTCCTGG GCCCAAGTGA TCCTCCCACC TCAGCCTCCC
10651 CAAATGCTGG GATTACAGGT GTGAGACACT GCACTCAGCT GCTGTTTGCA
10701 TAAATAATTA TGTTCATTGA CACCTAGAAT ATTAGTGCTA GAGGGAGTTG
10751 AGAGATATTT TAGTTTATGC CCCATGCTTT TTGCACATTT GAAAATGGTT
10801 CACAGGTACT AAGCAAACTG TTGACAGAGG TAGGCTTGGC GCCTGGGCCT
10851 CCTGACATAC CTGTAAACTG ATTTACGAGC TTATACCTGT ATAGCAAGAG
10901 GTTACAATGC TGGTATTAAG ATACTTCAGA GATTTTTTTT TTCTCCCGGC
10951 CCTCTAGTGA GTTTAATTGC CCCAGAGCTG GTTGGCGTCC TTGAATTCCT
11001 CTAGCTCATG AGTAAATGAA GCTCTCATAG ATTTTTAGCC AAGTGGCTCT
11051 GGCAATGAAG CTAGGCAGGA TCGTCTCTGG GATTTCCAGG TCCTTTGCTG
11101 GCATTTTGCC AGGTACTTCC CTTGTGAGAT AGCTTGGGGG TCCTTCCTAC
11151 ATTGCAATTG TTGAGAGAAA ATGCGATCTC CCGTGGATCT CTCTGGTGCC
11201 AGACTGGGGT GTTTCCAAAG GAGTACCCTG GCACTGGACC TAAGGAGAGC
11251 CTTCGGCGGA GCACCATCCT CTGGCAGGTG GTGCTGGGTG TCGGGGCAGG
11301 GTGGGGTGCT GTGGCAGCAG TTGGAGGTCC TGTCTCCTCT CAAGGTAGCT
11351 GAGATAGAGT GCCCAGGCTT AAGGTGGGCA TCCAGCCACA TGCCGGAGGA
11401 CAGTCTGACG GGCAAGTAGC TGTGCCAGTC TGCCAAGTGT CGGGAGGATT
11451 TTTGTCATTT TTTATATTAA TGTACCCTTT TTTTGTCACT TGGTTCTTGA
11501 AGAGCAGGAA GTTGACTCTT TCACTGTGCT GTAATACTCT CTCATAGCAA
11551 CTGGGACTCT GTAGAGTGGT TGTTTTCAGA TTCTGACAGG GGTCAGGAGA
11601 TAACGTTTTC TGCTTGGTAC TACCTAGCTG TTGCAGGGCA GGTTACTTCA
11651 TCTTTTAGCA TTGATTCTTC ATCTTTAAAG TAAGGGGCTT AGAGTGACCT
11701 GTGAAGGCTC ATTTAGCAAT GGTCTCTAGG ATTCTAGGGG CCTACATCGC
11751 TCCCAAAATG TGTCCTTATT GTGTATCTTT AAGAAGCCCT TGCTCTTCTC
11801 TTTTGTGTAG TATTAATAGT ATTCCTGAGT AAATCCACCC AGGGGACACC
11851 ACTCTCACCA CCCTCCCAAC ACATTGAAAG GACATTTTTT TCTCTCACCA
11901 TTTTAAAAAT GAGCACATCT ATAAAAATAA AAAGGAAGAA GAGTTGTGGA
11951 TGGGAGATGT TGAGACGAGG GCCAGGGTGA GCACCTTTCA GTTTCCTGGT
12001 CCTCCTCTGA GCTGCTTTCA GCTACCATTT CTCAGTACTC AGGTTGGCAG
12051 CAAGAAAGGA GTCCCAGGGT CAGGGTATAA GAGTTACTGG TGGCCTCCAG
12101 AGAGTATAGG ATCAGCCTGT GGTCACAGCA GAGAGAAAGA GAACGGCATG
12151 TGTGGCTCTG GGATTTGGTG GGAGTTTCAG CAGAATTGGA TGATCCAGAG
12201 GGGATTTCTG TTTTCTTTTT TTTTGTTGTT TGTTTGTTTT TTTTTTTGAG
12251 ATGGAGTTTC GCTCTTGTCG CCCAGGCTGG AGTGCAATGG CACGATCTCG
12301 GCTCACCACA ACCTCCGCCT CCCGGTTTCA AGCGATTCTC CTGCCTCAGC
12351 CTCCCGAGTA GCTGGGATTA CAGGCATGCA CCACCACGCC CGGCTAATTT
12401 TGTATTTTTA GTAGAGACGG GGTTTCTCCA TGTTACTCAG GCTGGTCTTG
12451 AACTCCGGAC CTCAGGTGAT CCGCCCGCCT TGGCCTCCCA AAGTGCTGGG
12501 GTTGCAGGCG TGAGCCACCA CACCCAGCCC AGAGAGGATT TCTTGAGTGA
12551 AATGTGTTCT CTATTGAAGG CAGAGGAAAA AGAGTATACG ATGAGAAATA
12601 CCCAGATTTC CATCCCCCCA AGAGCTTGTA CATATATAGA TATACGTGTG
12651 TATATTGTAT ACGTGTATAA TATAGTAACA TACACCGTGT ATATACGTAT
12701 ATATATATAT ATATATATGG GATGATTTAT ATATATATAT ATACAGCAGC
12751 ACGATTGAAC TATTGCACAA GGTCCAAGAC ATTATCTCAG AAAGGAGTAG
12801 ATAATCCTGA CCTAAGGAAT AGGGAATGCG GAATTCCAGG AAGCACTTCT
12851 CTTTCATTTT CCCCCACTCC TCCCAAGCAG TGCCTCACTT CTGCCTTGTC
12901 TAGCTGTACT CCGGAAAATT AAGAAATTTA TGAGTGTAGC ACCACGTATA
12951 CCAATGGGAA GGATGGGAGT CAGAAGTCAA GTGAACTCAG CCCGCCTCTG
13001 TGTACTTTGC ACTTTTCCAT TTCCCTTGGT ACCAGGCACT TTCATACTTA
13051 ATCCATAGTG GAGCTGTCAC AGTGAGCAAC TCTGACAATG ACAGCTTCTA
13101 CCCCAGAGGC CACCCCAAAC ATGGAGCTAA AGGCTCCAGC TGCAGGAGGT
13151 CTTAATGCTG GCCCTGTCCC CCCAGCTGCC ATGTCCACGC AGAGACTTCG
13201 GAATGAAGAC TACCACGACT ACAGCTCCAC GGACGTGAGC CCTGAGGAGA
13251 GCCCGTCGGA AGGCCTCAAC AACCTCTCCT CCCCGGGCTC CTACCAGCGC
13301 TTTGGTCAAA GCAATAGCAC AACGTGAGTA GCTGTTACCT TCTCCTCTCC
13351 TGGGTGGGAT TCGTGTTCCT AAGCCTCCCT TGGACTTATT TTTCCCCCCA
13401 ATTTCATCAG TCCTCCACTT TACAGATGAA GGTCAGCAGT GAAGAGATTG
13451 GGCGAGTGAC TGCGCTGAGA TTTGCCTTCC TGGGCTGCCA CTCTCTAGGC
13501 AGTTTCTTAC TCTTTTTTCC TTTCAGCTGT GTTGGCCCCC CAAGGCTGGT
13551 GCCAAGTGAG AGCTTGGACT TAAAAAAAGC TTCTACAGAG GACATTCTTT

FIGURE 3D

13601 TAATTTAAAA GTGTGTCATC TGTGCTAGAA CCCCAAATAA TTTCCAAGCA
13651 TAATCGGAAG CTTCCTTTGC AAAGTCTCCC CCCGAATTCT GCCCCATCAC
13701 CAAATCAGTA TTCATTTGAC TGAAGAAGTG GGAAGAGAGA AGAATTAACT
13751 TCTGCACTTA AAAAATTCAG GGTTGGTAGG AAAGGAAAGA TAGACTTTGC
13801 ATTCTCCAAA GAGGGCTTAA TCTCTTGTCT CCAGAAACTG GGACCCCAGA
13851 CTCATTTGGG CTGAGTTTGG CCCGCTTCAG GTCTCACTTT CCCCAAATGT
13901 AAAGAAAAAT TGAGGACTCC ACCACAAAGC TATGCTGGCT GTGTGGGGCT
13951 CACCACTTGA ATTAGAAAAT TCAGAGGAAG TTTTGCTACT CCATTGAGTT
14001 AGTTTCCCAG CTACTCCTGA TTTCAGCAGA CCTCTGACTT TTCTCTGTGT
14051 CCCAGCATCT CAGCTTTTGC AGTCCTGTTT ATTCCTCAAG CTTAGCTATT
14101 ACCTTTTCTG TGTTTTCTTG TGGACTGAGT GTGACTTACT GAGAGATCCT
14151 TCATGTCCTA GACTTATGCC ATTCCTGATG ACTGCCCAAG CGGACCATGG
14201 AAGCTTCTGG GCTCATCACT GGAGAAGCTC CCTCTGCCTG CACTGTCTGC
14251 TGGTACAGGG CATTTTCTCT TGCGAACTGG GGTGGAACTA GAAGAATGTC
14301 TGTCCACATT CCTGGCCCGT CACCACCACT AGCTGATTTC TATGCCTCAG
14351 GCTGGAAGTA CTCAACCAGT CCTCTAAGAT TCTGTTTCTG TAGCTTATTT
14401 CTCAGGGGTA TGCTTTTGTA GATTCCCCAT TAGCCTGCAG TGGGAGTTAG
14451 CTGGTGGTAG ATTGCTTAGA GCACAGCTGG CAGCAGTGTG GATCACCCTG
14501 CCCCTCTTTC CTCCAACCTT ATCAGCATTG GCAGCCCCCA TGCAGAAGCA
14551 TCTCCACACA CAGCCAATGG CATGTGATGG CTTCCCTTCA GAGGTCATGC
14601 TTGTTATCGT AAGATACTTC TAAGCTTCCT TCTCTGTAGT TTCCTTTGCA
14651 GTTTTTGCTC CTTTTTGATC TCAGATATCA ACTTGTCTAA GCAATATTTA
14701 GCAGATGAGG TCTGGATTTT TATGTTTATA GAGACATCTC TGAAGCTCAA
14751 AACCTACCAA CTAGCAACTT TAGGATAGTA GCTCATAGGT TTTGGACAAA
14801 ATTATGTCCT TGTTTCTTGG AAATCGAACA AATCAGAAGA TACCTTCCTC
14851 AGGCTTGTAT TGTGACATTT TCCAGGGTAT ACTTTGTTCC GAGTTTCCCT
14901 TCCTGCCTTG ATGTTGTGAT ACAGTGTAGG TGACCAGGGA AGCCTATCTG
14951 TAGTTGATGG CAGGTATTAC AGTCCCATCA CAGGTGGTAC AAGATAAAGT
15001 AATTTGCTGG GGCTTAGAGG ACTGGTTGAG TACTTCCAGC CTGGGGCATA
15051 GGATCCACGC AAGGATTTAT ATAGAAAACA TGCCAGGTAT GATTAAGGTA
15101 GAGGTTGATT TGGAGGACCT TCTTAACCTA AATTAATATT TTAATATGTC
15151 GGAAGTGTTA GAGACAAGTT TTTGAGCTGG GTTCCTTTTA TATTTCTGGT
15201 TTGCCCCACC CTTTTATCTA GTTTGCGCAA GGAACAAAAT ACATGGAAGT
15251 ACTTCTACAC CTACTGCACA TATGCATGCA CACACCTGGC TCTTCTAGCA
15301 AGTCAAGGGC TCAGCAAAAA CCCCTAGTTA GGGGGTGCAA ATAGGAACCC
15351 CAAACACTTC CATGAGTTTC ATGGGTTACT TCCTTTTATT TTTTTGAGAC
15401 AGGGTCTTGC TCTGTTGTCC AGGCTGGAGT GCACTGGCAC AATCATGGCT
15451 CACTGCAACC TCCATCTCCT GGGCTCAAGT GATCCTCCCA CCTTAGTTTC
15501 CTAAGTAGCT GAGACTACAG GCATGCTCCT GGCTACTTTT TGTATTTTTT
15551 TTTTTTTTTT TGTAGAGATA GGGTTTTGCT ATGTTGCCCA CTTAGTCTTA
15601 AACGCCTGGG CTCAAGTGAT CCGCCTGCCT CGGCCTCCCA AAGTGCTTGG
15651 ATTATAGGCA TGAGCACCAT GCCTGACCTG TGAATTATTT CTTAGTGTGT
15701 TCAGTGAGGT TATTTACTAA CACTTGATGT TACCAAGCTA TTGACTGCTT
15751 CGAAGACAGC CTCATTTTAT GCTGTTGGGC AGATTTTTCT TCTTGTTGCC
15801 CCTCTGAGTT CCATTATATA TATCAAGCCT CCGTGCTTCT TCCCCATGCA
15851 AACTGAAACC AGCAGACTGA AACTGGCTCT CTAAAGGTGA GCTGGAGTAG
15901 TCATTTGCAA AATGTGGTCT GCACACTTTG TGGGCTTCCC AAGACCATTT
15951 CAAGAAGTCT ATGAGGCTAA AACTCTCTTC ATAATAATAC TAAGATGTTA
16001 TCTGCTTTTT CACTTGTGGA TATTTGCACT TATAATGTAG AAGCAATGGT
16051 GGGTAAAATT ACACTGTAGA ACGAATCAAG GCAGTGGCAC CAAATTATAC
16101 TAGTTGTCGT TGTATTTTTC ACTGCCACAC ATGCGCAAAG AAAAAAGCCC
16151 TTTGCACTTA AGAATGTCTT TGATGAAACT GTAGGATTAC TAATATTTAA
16201 AAATTTGAGA CCCTTCAGTA TAGGTCTTTA ATATTCTGTG TGGCAAAATG
16251 GGAAGTATGC ATGAAGTACT TCTATGAGTA CCAAAATATG TTACTTGTCT
16301 TAAGGCAAAG ACCTCGAGTG ATTATATGAG TTGTCAACCA AACTTGCTGC
16351 CTTTTTTTTT TTTTCATAGA ACTAGAAAGA ACAACTAACA AACTGTAGGT
16401 CATTCAGACC CGAGTACTTG TAAGACATTT TCTTGAAAAT GAAAGAAATC
16451 AGCCCATCAC CTCAAGGAAA ACAATAGATA ATACATATCT GTTGCCCAGA
16501 ATAAAATTCA AGCTTTCAAG CAAAATTAGG AAAAAAAACC AACTTGTATC
16551 CAGTACCATG AGCTTGATAG CCCCTCTACT TGAAGACTTT TCTGATGAGA
16601 TTAGTGGTGA TATTAACAAA TATGACTTTT TGATATTATT AATATACAAT
16651 GAAGATGTTA ACATTTGGAA GATCTGTGTA AACTCAACCA AAGTATGATG
16701 TTAGGAATTC TGCATGGGTA AAAGATCCAT TGAAAGAGCA AGATCACCAA
16751 TGGATTTTTT TTTTCTTTTT TTTTTGAGAC AGTCTTGCTC TGTCACCCAG
16801 GCTGGAGTGC AGTGGCACAA TCTTGGCTCA CTGCAACCTC TGCCTCCCGG
16851 ATTCAAGCGA TTCTTCTGCC TCAGCCTCCC GAGTAGCTGG GATTACAGGT
16901 GCCTGCCACC ACTCCCAGCT AATTTTTATA TTTTTAGTAG AGACGGGGTT
16951 TCGCCATGTT GGCCAGGATA GTCTCAATCT CTTGACCTCA TGATCTGCCC

FIGURE 3E

17001 GCCTTGGCCT CCCAAAGTGC TGGGATTACA GGCATGAGCC ACTGCACCTG
17051 GCCTGACTTT TTTTTTTTTT TAAATACTAA ATGTATCAGG GACTTCTGGC
17101 CTCTTATGGT GTGGTGTGAC TTTTATGCTG TTCACTTTGT ATCTTTCTGT
17151 TACAGGGTTT GGGGCTTCTG TTATTATTAT TATTATTTTT TAATTTCCTC
17201 TGTTCTCTTA CCAGTGTTTG TCCGTCATTG TTTGGTTTGT CATCCTCTGT
17251 TGCAGTTTTG GGATCTGAGT CTTTTTTTTT TTTTGAGATG GAGTCTCCCT
17301 CTATTGCCTA GGCTGGAGTA CAGTGGCACG ATCTTAACTC ACTGCAACCT
17351 CTGCCTCCCG GGTTCAAGCA ATTCTCCTAC CTTAACCTCC TGAGAAGCTG
17401 GGATTACAGG CACATGCCGC TATGCCTGGC TAATTTCTGT ATTTTTAGTA
17451 GAGACGGGGT TTCGCCTTGT TGGCCAGGCT GGTCTCGAAC TCCTGACCTC
17501 AGGTGATCCA CCGCTTCGGC CTCCCAAAGT AGTGGGATTA TAGGCATGAG
17551 CCACTGTGCC TGGCCAGGTC TGAGCCTTTA CAGTGGTCAG TTCAGTGGTT
17601 AGAACCAGAC CCAAATACAC TTGGAAAGGA TAGAGTGTCT GAAGAGAGTT
17651 GGAGCACCCC TCTGGTCTAA TCTCTGAGAG AAGGGATTCT CAGAAATGTC
17701 AGAGAGTGGA GACTTACAGC ACAGTGGATA AGAGGGGGAG CTCTGGAGTC
17751 AGACTGCCCA AATTTGAATC CTGCCCCAGC CCTTTACTAG GTATGTGACC
17801 TTGAGCAAAC TGCTTCATCA TCTATAAGAT AAAATCTTAC AGGGTTGTTG
17851 TGGAAATAAA ATAAGATAAT GCATATAAGC ACTGAGATCC TAATAAAAGT
17901 TAACTGTCAT GGTTATCATT TCCTTGGCTG TCTTCCACTT CAGATGGTTC
17951 CAGACCTTGA TCCACCTGTT AAAAGGCAAC ATTGGCACAG GACTCCTGGG
18001 ACTCCCTCTG GCGGTGAAAA ATGCAGGCAT CGTGGTAAGG GTCTGCATCA
18051 GTGGAGAGGA GTGGTGACAA ATTTTAGGAG GTAGCTTTTT GTTGTTGTTA
18101 AAATGTACTT GCTTTAAAAC ATTTTAAATA GAGAAGCATT TTAAAAAAAT
18151 CAGTTGACAA AAAGCGGAAT TCAGACATTC ATTCACTTAA AGATATTTAT
18201 TGAGAGTGTT CTGTGCGTTA GGCACTGTTC TAAGCTCTTA GAATACATCA
18251 GTGAATTAAA TATTCCTGCC CTCATGGAGC TTACTTCATG GTGGAGAGGA
18301 TGTACTGAGA TGGCTCGAGC AGTTTCTGTC AATAATATGA ACTAATGAGT
18351 TAGTTACAGA TGTCTGCCCA TTTTCTACAG TCTCCCATGC CCTGTTCCTA
18401 AATGGCCAAC TGCAAGAATC TTATGTCTTC TTTTTGTGAT TTACCTCCAG
18451 TTGACTGCCT GCCCAAAGCC ATTCTGGTTT CTTTCGGAGT TGAAGAGAGA
18501 CTCAGAGATG TGGGTTGCCC TTAGCTAAGT GCAGTCTTTC TTGATCTGGC
18551 ATTGCTGTAA AGATAACTTA CCCGTCTCAC CTCACATCCC TTAGCCCAGC
18601 TCTTCCCACA GTCACAGGAG CCTTCTATTC TGCTGATGTG CACCAGTCTT
18651 GGAACNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNTGTTCT GTGCGTTAGG
19851 CACTGTTCTA AGCTCTTAGA ATACATCAGT GAATTAAATA TTCCTGCCCT
19901 CATGGAGCTT ACTTCATGGT GGAGAGGATG TACTGAGATG GCTCGAGCAG
19951 TTTCTGTCAA TAATATGAAT TAATGAGTTA GTTACAGTAT GTCTGCCCAT
20001 TTTCTACAGT CTCCCATGCC CTGTTCCTAA ATGGCCAACT GCAAGAATCT
20051 TATGTCTTCT TTTTGTGATT TACCTCCAGT TGACTGCCTG CCCAAAGCCA
20101 TTCTGGTTTC TTTCGGAGTT GAAGAGAGAC TCAGAGATGT GGGTTGCCCT
20151 TAGCTAAGTG CAGTCTTTCT TGATCTGGCA TTGCTGTAAA GATAACTTAC
20201 CCGTCTCACC TCACATCCCT TAGCCCAGCT CTTCCCACAG TCACAGGAGC
20251 CTTCTATTCT GCTGATGTGC ACCAGTCTTG GAACAGACTT ATCTTATGTC
20301 CTTCTTCTCC ACGTGACTAA ATCTCTCGAA AATGTGCTAA AGTTCAGATA
20351 ACACCCATCT CACAGAGCTA ATCTCCATGA TCACGTTTCT CTCTCTAATG

FIGURE 3F

20401 CTGGGCCCAA AGTTTTGTCA CTGAAAACTG CCTTAGTAGC TTTTTAATCC
20451 TTTGTGAACT GAGTATCCAT TGGGTTCACT CCTAATTCTA CCTACTTTTC
20501 TCTCTCTCTT TTGCCTGCAA TATCTGTCCC CAGATGGGTC CCATCAGCCT
20551 GCTGATCATA GGCATCGTGG CCGTGCACTG CATGGGTATC CTGGTGAAAT
20601 GTGCTCACCA CTTCTGCCGC AGGTGAGAGC CCTCTGAGCC ACCTCTCAAG
20651 TGACAGATTG TCCTTTTGGG TTCTGTTATC AACCCTGAAA ATGAGCACTG
20701 ATGCAGACCA CTCTCAATTC TTTACACTGG CTGGAGGTAG CAGCTTATGA
20751 TTGCAGCGTT TTCCTTTCCC TGGTTATTTT TGCGTTCTTT TCTGGCTCAT
20801 TATCATCTGT TAAATTTACT TATGCCCAGT GGGTACTACA TTCTAATTTC
20851 ATGGGCGTTG TAATATTTAC CCCATTGAAA TGATTCTACC AGATGGTTGT
20901 TAATTATAAT AAAAGTAACC ATCCTGTCGA CTGAATACTT CTGATCTTTG
20951 AAAGCACGAG ATACAGGACT CAGAGTGGTA CCTCCAGGGT GAAAGATGGG
21001 AACTGGCCCA GGTCTCAGTG GCTCTTTTTG TTCTGTCATT GTCATTGTCT
21051 AATCCACGTG CTCTGTCCTT CCTCTTCCCT CCTACTCTTC CAGGCTGAAT
21101 AAATCCTTTG TGGATTATGG TGATACTGTG ATGTATGGAC TAGAATCCAG
21151 CCCCTGCTCC TGGCTCCGGA ACCACGCACA CTGGGGAAGG TAACTGATTT
21201 CCTCCTTCCT TTCAACTGTG GCCTCCCAGT GTGAGGCCTT CAGATGGGGA
21251 GGTGCAACGT GGGAGACAGT GTAAAGCGTG GAAAGAGTGC TGTTTGGGTC
21301 AGTTGCCTTG GGCTGTGGCT CAGCTCTGCT GGTAGTAAGC TGTGTGACCT
21351 GGGGCTGGGT AACCCCTTTT TTCCTTGGGT TTTAGTTTTC TTATCAGGAA
21401 AGCATGGGGC CTGGCCTGAA TGGTCTCTAG AGCCATTCCA GCTTTGGCGG
21451 TCTATGACCA GTGATTGTTT TTGATTCACT CATTTGTTCA ACAAATGTAT
21501 TTAAGCACTA TCTTATAAAT GGAACAAAAC AGTTCTAGGT AAGAAGGGAA
21551 GATTTCCTGA AGTAAATTAT GTGGTTCCTA CCCTCCAGAG GCTTGTAGTC
21601 TGTGTAAGGA AAAAGAAATG TGGGAAGAGA AGCCGGGGAA CAAGATAAGA
21651 GACCAGTAGT GGGAGACACC CATAAGAAGA AAGTGTCATG AGCTAGGAGT
21701 ACACCCTCAG TGCTCAGAGA GAGAGGAACT TTAAAGATTC TCTTGTCGGC
21751 TGTGCCAGAT GAGAAACGCA CATGAGAGAT AGGAGCAAAG AAGGCTTCAG
21801 GAGAAGGTGA GATAAACTAG AGCAGGGCGT GGAGATGAGT TTGGAGGTGG
21851 GAAGTATTTG CAAATTTCTC GTTATGGTAA CTCTTCAGTG TTTGGAGGGA
21901 AATATTATGT TTGTTTTCTA CATTTAAATG TAGGAAATTG ATACTATCAA
21951 GGGCTAAAAA TTCTTAAAAA AAAAAAAAAA GAACCACATT AAAACTATGT
22001 TCTCTAGAAA AGTTCCTTTT TGTTGTCATA GAGGAAACTT ACTTTCATTC
22051 ATAGTCACCT TTATCCTGTG ATGCAGATTA TATAGTTCTT TTGGCCAAAT
22101 TATTTTCTGT AACTGGGAGA AGCTAGATTG CCAGGTGACC ACCATGAGTT
22151 GGGTGGTTGT TAATTCTTCC TTCCATTCTT TCTTACTACT TCCTTTCTTC
22201 CGCCCTCCTT CCCTCCTTTC CTTCCTTCCT TTTAATAAAA TGTGTGCTAT
22251 TTTAATGCGT GCTCATAGTA AAAACTTTGT TTTGATCAAG ATAGGACATA
22301 AAGTAAAAAG TGAAAGAAAA TTTTGGTCAC AGTTGCATGG GTAGCCTTTT
22351 GGAATTTGCT GTATAAGTAG AAACATACAC ATGTTCTTAA AGTTTTTTGC
22401 ACAGATTGAC CATACTATGT ATACTGTTTG GAAACTTGCT TTTTCCCCTT
22451 AAACGTCTGA GACGTTTTTC TCTATCAGCA CATAGAGATT TAACACATTC
22501 TTTTTAACTG CTGTGTAATG TTCCATTTAA GAACGGTCTA TAATTTAATC
22551 ACTCTGCTTT TGATGATCCT TTAGGTTGTT ACCAGCTGCT ATTGTTCAAC
22601 CAGCAGTCTG TTTTTGGTAC ATCAGTTTCT GTGTCCTTAA TGTGGGACTT
22651 GGTTGGTTCT TATATCCAAG TTATAGAGAC AGTGAAGGGG ACTATTTCCT
22701 TGTGTTTTAT GTCAAGGGCT CCCTGTAACT AACAAAAAAG TGTGAGATGG
22751 GATAGGTGGG CAGATGTGTA GAGAGGATGC TAAGGGGCTG GGCAGTGGTC
22801 ATGGTGTCTG TGCATGTGTC TCACCTCATG CAGCATTCCA GACGAGAAGC
22851 CAGGAAGGGG ACGTCGGAAA CCACACAGAT AGCACCTCCC TCACCTTCTT
22901 CCCAATGCCC CAGACCAGTG GCACCTAGCA TGGTTTCTTC TCCTGCCAGG
22951 GCATCTCGTC CTTGTCACTG CCAGGAAGGG TCTGTGATGG CTTGGGGAAA
23001 AGCACTGTTA AAAAAACACT TAATGGGCAC AATGTACACT GTTTGGGTGA
23051 TGGGTACACT AAACGCCCAG GCACTACCAC TATGCAGTAT ATCCATTTAA
23101 CAAAACAGCA CTTGTACTCC CTAAATCTAT TAAAAAACAA AAACAAAAAA
23151 CACCTCCCCT TCTGGGAGCA TTGCATTTGT ATTGTAACAG TCTTTGTATT
23201 CCTTCCTTCC CCACCTCCAG ACGTGTTGTG GACTTCTTCC TGATTGTCAC
23251 CCAGCTGGGA TTCTGCTGTG TCTATTTTGT GTTTCTGGCT GACAACTTTA
23301 AACAGGTAGG CACCTGGTTA AAAAAGAAAA AAAAAAAAAA AACCAGAGCG
23351 AGAATGGCAA AAGATGATTG AAGTTTTTGT TTAGGATTTT TTCCAAATCA
23401 GCTTTTGTCA ACAAAAGAGT TAAAGTTTTC ATATTTTACA TAGATCTACG
23451 TCTTCTATTT GATTCCCATG GAAAGAGCTC GGGCATAGAG AAACCGCCAC
23501 ATGTCTTGTC GACCCTCCTG TCCTAGGTAC ATATGATCAA ACCTAGCTCA
23551 GACAATTGGG TTGCTGATGA TAGTCGTGAA GTTCTCTAAA GATGGCTCAC
23601 TGGCCACAGA TTCTAAAAGG CCTTGTTCAC ACACCTGAGC CTTTCCTCAG
23651 GAACCTCTTC CAGCAGAGGA TCCACCGGCC TCTGTTGTTT GAGAGGTGTT
23701 TCCGTTTTCT TCCTTCCCCT CATTCTAGGT GATAGAAGCG GCCAATGGGA
23751 CCACCAATAA CTGCCACAAC AATGAGACGG TGATTCTGAC GCCTACCATG

FIGURE 3G

23801 GACTCGCGAC TCTACATGCT CTCCTTCCTG CCCTTCCTGG TGCTGCTGGT
23851 TTTCATCAGG AACCTCCGAG CCCTGTCCAT CTTCTCCCTG TTGGCCAACA
23901 TCACCATGCT GGTCAGCTTG GTCATGATCT ACCAGTTCAT TGTTCAGGTA
23951 CATGCCTAGG CCCTCTCCTA TCATCTTGGT TCAATATTTT AAAAAAGCCA
24001 GGCGTGGTAG CTCACGCCTG TAATCCCAGC ACTTTGGGAG GTGGGGGCGG
24051 GTGGATCACC TGAGGTCAGG AGTTTGAGAC CAGCCTGGCC GTCATGGTGA
24101 AACCTGTCTC TACTACTAAA AATAGAAAAA TTAGGCATGG GGGTGTGGGC
24151 CTAATCTCAG CTATTTGGGA GGCTGAGGCA GGAGAATCGC TTGAACCTGG
24201 GAGGCGAAGG TTGCAGTGAG CTGAGATCAT GCTACTGCAC TCCAACCTGG
24251 GCAACAGAGC AAGACTCTGT CTCAAAAAAA AAAAAAAAAA AAAAAAAAAA
24301 TATATATATA TATATATATA TATATATATA TATATAAAAT ATATATTTAC
24351 ATATATGTGT ATATGTTATT ATATTTTACA TATATATTAC ATGTATATTT
24401 TACATATACA TGTAATATAT ATTATGTACA TGCATAATAT ATATTATATA
24451 ATGTATGTAA TATTTATATA TTATGTATAT ATATACATAA TATATATATG
24501 TAAGTGGAAT GTAAATAGTT ATATGTTACT ACTGGTATGT CTAGATTAGA
24551 GGTTCTGTTC TTGGGCCCTG TTGACATTTT GGGATGGATA AATTCTTTGC
24601 TATGGGGCTG TCCTGTGCAT TGTGGGGTGT TTAGCAGCAT CTCTGGTCTC
24651 TCTCATTAGG TACCAGTAGC GATCCCTCCA TGAGTTATGA CAACCAAAAA
24701 TGTCTCCAGA CATTGCCAAA CCTTCCTGGG GGGCAAAATC GCCCCCCCAC
24751 CCAGGGGGCA CTGGTTTAGA CTTTTTTCAA TTAGATGGTT AATTCATGAT
24801 CATTGTATAC AGTTGGAAAA TAGAGGAAAA TGTTAAGATT AAAATAAAAA
24851 ATAATTTTTC TAACCTGTAT TTAGATAAGT AATTCCTTAT CAACTCCAGT
24901 TAATTTTTAT TTGTCAAAAT TATAAATTCA CTTGTTCCTT GCCCTCACTT
24951 AACCCATGCA GGCAAGTCTG TGGGGTGGCA TGAGAGAGAA CATCTGTATA
25001 CAGATGGGTA GAAAATCAGG CTGAGAAAAA TGTGCCCTTA AACACTATGG
25051 CTGTTTGTGA AAATGAGAAT GCATTTTCTA AGGCTTGAGA AAAGGAAAAA
25101 AGTAAAAGCG GGTAAATAAA AGCATAACTT AAAAAAAAAA ATACTTAAAT
25151 TCAGTTCCCC AAATAATTCA TCAGTACATA TTCATTAAAA TGCAGACAAC
25201 ACAAATACCT CTTGAATACC ATGTCCCCAC CCCGAGTCTC CTCTCAGGGA
25251 CCCGCTGTAT GTGATTGGTC TGTCTCATTC TAGATCCTGT GAATGGATTT
25301 ACAGCCCATG TAAGTATATT GAGAAATACA TTGAAATATA TTTTGTTTTC
25351 ATTTTTGAAA CATAATTTTT TAAAGTTACA TGTTCATCTA CCTTGCTTTT
25401 TTCCACCTTA AAAATGCCTT AGTGAGCCTT CCAGGTTAGT ATTCCTGGCT
25451 CTACCTTGTT GTTGTTAGTT GTCACATTGT ATCACAGCAA GGAGATTTGC
25501 TGCCATTTAT TTAACAAGTC CTCACTCAGT GGCTATCAGG CCATGGATAA
25551 TTTTTAGTAT TATTTCAGTA TTAAGACAGT GAGATGTTCT TATACATTCC
25601 TTTTTGTGGA CTTGTATAAA TACTTAAGAT ATTTGTCTAG ATGTGTAATT
25651 GCTGAAGAGT GTGCACTTTT GAATTTTTGT TATGTTACCA AGTTGTTTTT
25701 TCCAAAAGCT ATTTCCTAAT TTCATAGAGA CTCCCTTCCA CAGTATTTAA
25751 GTGCCCATTT TTCCATCCTT ACTAATACTG GATGTATTAG TATTATTTAT
25801 ATTAGTATTC TTGATATATG TATGATATAT TAGTATTTAT ATATAGTATA
25851 ATTATATTAG TATTATTTTG GTCAGTCTGA CAGGTGAATA TTATCTCATT
25901 TTCATACAGT CTGCTTAACA GTGACCCAGT CACCCACTGA TATAGTTTCC
25951 AGGAAGACAG TGGCTCATAA AAAGCAGGAC TTCTTGTGCT AAGCAAATGA
26001 CATTATCAAT TTAGATTAAC ATTTTGCTCT GTGAGTATTG ACTGTTTTTT
26051 CATCACATTA AGTCAGATGA GTGGCAGATA TTGACTCTTC TGCAGAACTA
26101 TTTTTTTAAA CATAAAATAA ATAGTTCTTC ACGTCCCCTT TTTATGAGAC
26151 TGGAAGGAGG GCCTGTGAAT GCACTATAGA CCATTTCCAT TACCAAACAA
26201 GGAAGGGTAA ATATTTACCC ATGATTTCAT GTTGTAATTA GAACTCTCAA
26251 GAGCACCTAG GATATACTCC ACTGTACACT AAAGATGAAT TGAGTGAGGC
26301 CACCCCCATA GACAGACACA CCTGCTGCAT TTACTTTTAA ATCATTGCAG
26351 GTAGTCGGGC TGTGGTGGTT GCCATGATCA GAGGGCTGGG ATAAGAATTT
26401 GGGTTCTTAT AGCGTCTCAG TCCCAACCAA CTGGTAGTAT CCATCCAGAG
26451 TGATGTCTAT GCATAGTACA ACCAGGACAC AGAGCAATGT CTGCATAAGG
26501 GCAGCCCTGC TGATTTCTTG AGAGCAATTC TGAGTCTTCC TCTGGGCTTA
26551 GCCAGAAGTT GTGCTGTGAT CAAATAGTGC CGTCTGCCTG GAGTACAGCA
26601 TGGGGGAAGA GGTTTGGCTG TGTTTTGATG TAGTCACTGC CCATAGTGTT
26651 GTAGTTGCTT CATTTTGATG TGTCATACAG CTAAAGATGC TCCCTTTAGG
26701 TCATTTTTGT TGCCGCTGCC TCTGCGGCTT GTTACTACTG TTCTGTTTTG
26751 GCATTGTGCC CCACTTACCA TGAGGATTCC CCTACTGTTC AATGTTTCTG
26801 AATTTTTTCC CTAATCCTAA GCATGTACAT GACTGTTCCT CTTGCCCCTC
26851 ATGCACGTGC CATTGTAGGT AGCAGACCAA GGTCTTCCAC AGAGAGCAGG
26901 TTCCTCTCTG TCTTCAGCAT GTGGAGTCTC AAATGGAACA GTTCTGGGCA
26951 GAGTGCTTTG CACAGAGGGT GCTCCCAATA AATGTTTTAT CACTGCATAT
27001 CGTTGCTTCT GAGATGTATT TTTTCATAGT TATAACAGTT TCAGGATTGC
27051 AAGAGTACAT CTCACAATCC ATGTGTACCT TTAACAGCAT TTTCTCAAAA
27101 TACTGTTATT ATAATTGATA ATATGGTAAG ACCTCACTTA ATATCATTGA
27151 TACATTCTTA GAAACTGCAA TACATTAAAT GTATGTATAG CGAAATCAGT

FIGURE 3H

27201 TTTTTTCTCA TCAATGTTAT AACAAAACAG CGTTGAAGGA AGTGACTGTA
27251 CGTCATTTCA CTTAAAGTCT CAGTTTCCAA GAACTTATTG ACGACAAGGG
27301 AGGACTTACT GTGTTGTAGA ATTGAGGAGA TATGTTAATA ACGAGCTGAT
27351 TTTAACATGT ATGTTTCTTT ATAAATTAAA CTTTTCTCAT TTAGTTGGTT
27401 GGGTCAGTAG CAATCAGTAA GTATGTAGAA TAATACACTT CTTCTGCTGG
27451 CCTCATTCCC ACAATATCCC CACATATGGA TTGTGAAATT CCCAGTCTGA
27501 TACTTGAATC TGATCTGATG TATGAATAAG AGCAGGAGTC ATTCACTAAC
27551 CAACAGATAG CACCTGTTTC CAATAACTTA GGTTACATTT GTGACTCAGG
27601 AATAATTACA GGCCACTCTT GCTCTCAAGT CCCATTGTAA AGGAAAAATA
27651 CCTATTACCC TGTCTTCATT CCAGGTATTG AAATGCTTCT TACAAAGGGA
27701 TCTAACAGAT TTCTTAGCAG GGGCCCAGGG AAACACATTT ATTTAATTTT
27751 TTTATTTTTT CAAAAGCAAT ATTACTGCTT TGAAATCTTT CAAAGTGAAG
27801 GCTGTTATAG AGCTTAATAA TGGATCTCCT TTTACTTGCC TGAAATTATT
27851 CTGAAGCCTG TTAAGAGCAT GCCCCGTATT ATCCAAATAG CCATACAGTT
27901 AAATCAATTT TAAAACATTG TAAAAGGCTG TTTTAACATC AATTTTTATT
27951 TTAATTGAAG CAACATACAC ATGTGGTTTA GAAAACCAAA TTGTAAAAAG
28001 ACAGCAGCTT TGAATCCCTC CTCCCCACCC TGCCCCTTCC ACACAGTCTG
28051 TTACTGGAGA CTGTTGTTCG GTGGAGGATT TTGTGACTAT ACCTCTGTCT
28101 TAGTCAGGGT TCTCTAGAGG GACAGAACTA ATAGGATAGA TGTACATATA
28151 TAGGGGAGTT TATTAAGGAA TATTAACTCA CGCAATCACA GGTTTCCACA
28201 ACAGGCCGTC TGCAGGCTGA GGAGCAAGGA AGCCAGTCCA AGTCCCAAAG
28251 CTGAAGAACT CGAGGTCTGA TGTTCAAGGG CAGGAAGTAT CCAGCACGGG
28301 AGAAAGATGT AGCCTGGGAG GCTAAGCCAG TCTAGCCTTT TCACATTCTT
28351 CTGCCTGCTT TTTAATCTGG CCACTGGCCG AGCTGGCAGC TGATTAGATT
28401 GTGCCCATCC AGATTAAGGG TGGGTCTGCC TTTCCCAGTC CACTGACTCA
28451 AATGTGGCAA CACCCTCACA AACACACGCA GGAACAATAC TTTGCCTCCT
28501 TCAGTGCAAT CAAGTTGACA CTCAGTATTA ACCATCACAA CCTCCTTCCT
28551 TATACAACCT TAACATTGTA CCTGCAGTTA ACAGTTGCCC TTTTTCTGGC
28601 CCATTTTTTA AAGCATTCTT TTTGCTCCTC CTCCCCACAT GTTCCAGCAC
28651 TCCTGTTGTG TGCTTCTTGG TAATACTTTG AAAGTGCTCA AGTTCATTGA
28701 TGAGAATTTT AAAAAGGAGA AGAAAAGAAG AGGAAAAAGG AAGAGAACCA
28751 ATATAAAAAT GTACCACTTT CTCTTCCCTT CCAGCTTTAT CTTTGATGTT
28801 TATGTAGTTG TATCAGAGTG AATATATAAA TTAAAATTAA AATTTTTTCT
28851 CACTAATATT TCGTAAGTAG TTTTTCATGT TCTACTTAGT TATCTCAATT
28901 TTTACTTTTT AATAGTGCAT ACTGGCCAGG CGCAGTGGCC AACGCCTGTG
28951 ATCCCAGCAC CTTGGGAGGC CGAGGCAGGC AGATCACTTG AAATCAAGAG
29001 TTCACAAACA GCTTGGCCAA CATGGTGAAG CCCTGTCTCT ACCAAAAATA
29051 TACAAAAAAT AGCCGAGTGT GGTGGCACGC ACCTGTAGTC CTAGCCACTC
29101 AGGAGGCTGA GGCACGAAGA ATTGCTTGAA CCCAGGAGGC AGAGGGAGGT
29151 TGCAGTGAGC CGAGATCTCG CCACTGCACG CCAGCCTGGG TGACAGAAGG
29201 AGACTCTGTC TCTAGATAGA TAGATAGATA GATAGATAGA TAGATAGATA
29251 GATAGATAGA ATATATATTC TGTGGTTTAG CTGGTATGTT GTTAATTACT
29301 TAACTGATCC CTTGTTTGGA AGCACTTATA TTGTTTTCAG TTACTTTATT
29351 AAACAGCTTT GCTCAGTGTT TTTCATCTTT TGATTTTTTT TCTACTTGAA
29401 TTCAGTTTCT GGGGATGGGA TTACCTAGTG AAAGAATGTG ACTCTTTTTA
29451 TGCAAGCCCC AACATTTGAG TTTTTAATAG TACCTGGGGC TTGTCTTTCC
29501 CCCAAACAAG TGGGTTTTTC TTAGCCTGAA GAGAAAAACA TACAAAGGTT
29551 AAATGTCCCT AAATCATCTG TCAGGTATTA GACTTTCTTC CTTTAGAGAA
29601 TCTTGGATTT GTTAAAAGGT ATGACCTCTC CGATTCAGAG TTCAAATCTT
29651 GAATTTCTGT ATAGCCTTTT GCTTTGTTTT GCTTTCTGTC TTTCAGAGGA
29701 TCCCAGACCC CAGCCACCTC CCCTTGGTGG CCCCTTGGAA GACCTACCCT
29751 CTCTTCTTTG GCACAGCGAT TTTTTCATTT GAAGGCATTG GAATGGTAAG
29801 AGCTGCACTG TGATTTGGGC TAGTGTTCTC TGGTGCCCTT GGTGTTCTCC
29851 AGGTCTGTTT CAAGGAATGC TGAGGAAACA TTGTTAGAAA GTATCTTCTG
29901 AGGCCAGGCA TGGTGGCTCA CGCCTGTAAT CTCAGCACTT TGGGAGGCCT
29951 AGACTGGTGG ATCACTTGAG GTCAGGAGTT CGAAACCAGC CTGGCCAACA
30001 TGGTGAAACC CCATCTCTAC TAAATATACA AAAATCAGCT AGGCATGGTG
30051 GCACACGCCT ATAATCCCAG CCACTCGAGA GGCTGAGGCA GGAGAATTGC
30101 TTGAACTGGG GAGACGGAGG TTGCAGTGAG CCAAGATCAC GCCACTGCAC
30151 TCCAGCCTGG GTGACAGAGC GAGACTCTGT CTCAAAAAAA AAAAAAAAAA
30201 AAAGAAATTA TCTTCTGTAA CTCACTGGTC AGTTAGTGAA TAGTGTTTCG
30251 GGGATTCCAT TGAGATTTCC CAGCTTCAAC TTTTCAAGAC AAATTATATG
30301 TAATTTTAAA ATGTTTACAT TCAAGGCCCC TTCACTGCAC ACTCATCTCC
30351 TATGTGTGCA GTAAGGAATA GCATATGGCA ATCAGGAAGG CAGGGTCTAG
30401 AGTCAGACTG ACATGGGGGT AAGTCCTGGC TCTGCCATAG AGTAGCTGTG
30451 TGACCTTGAG CAAGGGCTTC ATCTCTTTGA GCCTTCATTA TTTGTTCCTG
30501 AAAAGTGAGC TTAATGATTC GTAGTTATTA GGATTAAATG AGATATGTGC
30551 AAAATGCTTT GCACAGACCC TGACACATGG TAAATGTTTA ATAGATTTTT

FIGURE 3I

30601 ATTTTATTAA TAATGTTATT TTATTATTGA ATCAATAAAT GCATGAATAA
30651 TTTCTCTGCC CTACAACATT GGTTTGGTGT ATTTTCTGCG TGCAAAAGAG
30701 CAGCCTTCAC TCCTGGCTCA GCATTCTGTG ATTTCACCAA ATGCTTTTCC
30751 TAAAAAGGAA TCTACCCCTC ACTTTTACCT AATTTGAATT TTATTTGTAT
30801 TTTTCATAAT AATGGTACAA AACTCTTCTC TGAAGAAAGT TATTCCTGGC
30851 CAAGGGCGCC ACAAATGGGA AGAGCCTCCT CCCGGCACGG CACTTCTTTC
30901 TCTTTTTTGA TTTCCGAGAC ACCTTATTTG CTTTTAAGAA AACCTAGAAG
30951 CTGTACACAT TTTTTGTCAA AATCGTGAGA AATCACCCAG GTGGTGTTGA
31001 TGGAACACGT TGAAGCTCTG ACATCATGGG GGAGGCTTTG GGAGTGATCA
31051 CATGTAAATC ACTGGTCTCC CTGAGGTTTT ATACCTTGCC CTGTGCTCTA
31101 TCTTAGGGCT TTTCATTGCC CATGAAGAGT GTCTCACTGT AATTCAGAAA
31151 CACAAATGGT TTCCTGCTGC TGGGGCAGAG GCCTGAGTGG GCCCATACTT
31201 CAGCAGTGAG AAAGAGATCC CAAGAACTCA GGACTGGAAA GAAGAGGCTG
31251 AGAAAGTGTG GAAAGATGCC CAGAGACCTT AGGTTCTTGG GCATCCTAAG
31301 GGACCTTGTG CTAAATTTTT AGTAGCTTTC CCTAACAGCA CAGCGCAGAA
31351 ATTGTTTGCT TGGTTTTATT ACCCAAGACT TGTACACAAA GTTATTCTGC
31401 AAACATCATT TGTTTTCAAG ATTTCTTTGT ATTTCTATTT TTTTACAATA
31451 GAGAGAGAAC ACTGCTAGAT TGACTCTTAG TTTTGGATCT AGGGCTTGTT
31501 CATTGCATCG GGGTAAAGTG CCAGGCTGCA CACTGTATTC ACCGTGTGCT
31551 CTGTGTTCAT GCAGCTGTCA CAGGCCAGAT ATGGGCTCCC TGCCCTCTGG
31601 CTCTTTGTAT TCTTGGTATG ATGGAAACTG ACAGATACAT TTAGAAGCTG
31651 TTTCTAATGG ATGTGGATTA ACTGAAGCTG GAGCAGGTGG TGGGGGACAC
31701 TGGGGGCCCC TGGGAGCACC AGGCCTGCAG CAGGATCATG TTGGTGTGTG
31751 CAGCACAGAC CTATTGTGCC TCATGGCTGC AGGTGCTGCA GGAAAATAAG
31801 TGTAATAACA AGATCAAAGT TTGTGGCACA GCTAGACATT GCCCAGTGTT
31851 GCTCTTCCTT ATTAGCTGTG CCAGATCCAG TTGATGGCAG ACTGTGAAGA
31901 TCCTCACAGT CACACCTGCC ATTTTCCCAA TTCCTAAATA AACTCATTTT
31951 CATAGGGGCC TTTCCTTTGC TTTTACCAAA GTTAAAGAAT GTCCTCTCTT
32001 ATGATGAAGG AGGTAGCAAA GGTCAGTGGT TTGTAATAGG AGTCTGGAAA
32051 CTGGGCATTG TTAAGCACCC ATGTCTTGAG ATCCTCTACG AAGTCATGCT
32101 TTCTTTCGCA CTGCAGTTTC TCTCACTTGG AGGTTTAACT TGTCACTGCT
32151 AGTGCTTGCC CCTGTTGAGC TAGAGTGGCA GTTTTCCCTG ACCTATATTT
32201 TGCATTCTTT AGACAGGTTA GCAGGAAGCT GTGATCTCAG GTTAGATGCC
32251 AGGTGGGCAT GACATGAGAG GGCCTTCTGG TTGCCATGTG GCCTCACTCA
32301 CAGGGCAAGG GACATTCCCA GCCTGCAGGG ATCCTGCAGC AGGAGGACAG
32351 AGCACTGGCC TGAGCCAGGA GTCCTGGGCT CGTGTCCTGT CCACCCTTAC
32401 TTTTAGGACT TCCTAGCTAG GCAGTGGGCT GCAGAGTCCC TTCTAGTCCC
32451 AAGAGCATAA CGTCTGATGA AATAACTTTA TTTAAAGAGC AGATGTGCTT
32501 CTGGAGAATT CTGGGGATAA AAGAGTTACT TTTTTTCTGA GGTTTTTTTT
32551 TTTCTTGGCC ATTAACTTTT CTTTTTTCTG CACTTTCTCT CCTCTCTCAC
32601 TACTCTCTCA TAGGTTCTGC CCCTGGAAAA CAAAATGAAG GATCCTCGGA
32651 AGTTCCCACT CATCCTGTAC CTGGGCATGG TCATCGTCAC CATCCTCTAC
32701 ATCAGCCTGG GGTGTCTGGG GTACCTGCAA TTTGGAGCTA ATATCCAAGG
32751 CAGCATAACC CTCAACCTGC CCAACTGCTG GTACGTGGAG GGAGGATGGA
32801 AACCTAGGAG CACTGGATAT TTTTAAAAAC TAATGGGTCA CAGTGTGGAT
32851 TCTCCCTCTT ACTTATCTCT TAAACCAGCC CACTTCACTC TAGCCCACCA
32901 TCCCCTGCCA CTGCCAGCCC TCACTGGCTG CCCTGGACTG CATTCTGTTT
32951 GGGGAATTCA TGTAGAGCCT TCTGCTGAAG CCATTGGTGC TGATCAGCCG
33001 ATGGGTAAGC CATTTCTCCT TGGAATTCCT AAGCTCAGAA GGACCGAGTA
33051 TCTAGTCCAT TCATGGTAAA CCATTCCAAA TAGACAGGGA GATGGGAGGG
33101 CAAACCTGCA TTTGATTCCC AGCATCGGTT GTGCCTCTCC CTTGGTAGTA
33151 ACAGGCTTGA TATGCAGATG GGAGCATCTC ACTGTGAGCC GGGGATTGTT
33201 GGGAGTCCTT TTGTACCTCC CTTGCATTGG TGAATGTATT ATAGGGAAAT
33251 AGTGAGCCAT TTTGAAATGC TTCCTGAAAG GGTGAATGTC CCAGGGCATG
33301 TGCAGAGCAA CCATCCTGTT TTGAAGATGA ATCATCTCAT GGTGGAGAGC
33351 AGCTGTTAGC AGACACTGAG AAGCTTGTTG AGTGCTCTGC GGATCAGAAT
33401 CAGCTTTCAG TCTAGGCTGG CTGATCTGCC TGGGTGTGCT TTTTATTTTG
33451 TTTTGTATTG TTTTATTTTA TTGTATTTTT TAAGACAACA GCACTCAGTA
33501 TTTCCAGGGG CTTTCCCGTT CAAGTACGAA CCAGGCTTGA CCCTGCTTAG
33551 CTTCCAAGAT CAGGTGAAAT TGAGCACATT CAGAATGGTA TGGCTATAGA
33601 CCTGGATTCG CTTTTTATTT TTTTATATTC TTTTTCAGTT GATTTTAACT
33651 CGTGAGGCAT ACCAATTATA TATATGGATG CAGTATGTGT GACATTTGGA
33701 TACATATGTA CAATGTGTAA TTATCAAATC AGGGTAATTG GCATATCCAT
33751 CTTGTGTCTA CTTTTAAATT TCCAAATGTT TCTGCCCTTC CAAGAAGGAA
33801 GAGGCAGGTG GTAGCTTGGT GTAACTGTGT CACCTTTCCC TGGAAGATAA
33851 ATGGATCGGG AGCAACAGAA GCAGCCCACA TGATCCGAAG CCATAGAGGA
33901 GAATCTGTCT TCTTTCCTAA CACCCCAAAC CCAGCTGCTG TAACTCTTCT
33951 GCCTCCATTT GGGTATAATT TATTTGGCTA TCCCTGCAGG TGTCACTCTC

FIGURE 3J

34001 CTAAGTCCAG ACTTCACAGC TCTCCAGAGG CTTTGGGGCT GCTTTGAGTT
34051 TAATGATAGA GCCACCAGAT GATTTTTCCC AAGAGTTTTT ATTATCTATT
34101 CATGGAGCAA GTATGACCTT TTACCAGACT CAGTCTTTAC AAGGTTGTTC
34151 TCCTGCTTAT AGCATAAGAA CATCTTCTAG ATTTTAAATT CAACCACAGA
34201 GAAACTCAAG GCACATATAC ACAGTCTGTA TTAGCACATT TAAATAGATT
34251 TCCGACAAGG GAGGACAAAT GTTTCTTGCT GTTTAACACA TGAGGGTCTG
34301 GTTTAAGGTG GAGCTTTGCT TAGGGACAGA GACCTTTCCT TTTAATGACC
34351 AGGTCAGATC TGTAAGTTGA TCACAGACTG TTTTCCTACT CTGTGCAGTC
34401 AAGGCACTGG AGTAATAAAA TAGGGATATC CTGTGGTGAG TTACGTCATT
34451 TTTGGAAGCT ACACTTGAAG CAGTAGTAGG AAGAGAGCCA TAGTGGTATG
34501 GAAAGATGGA ATTCTGCTCT GGCCTCTTGG TCCTGCAGTG TCTTCATCTA
34551 ATTCTAGGGA CACTGACTTG GATGGGACAG ATATAAATAG GCTTGTGACA
34601 TTTTAATTGC AATTTTGTTT TTATTTTTGA AGGCATGTAC ACCTGTATGC
34651 CCATGGCAAA GATTGAGATT TTCAAAAGGT ATATAGAGAG CATTAAGCTT
34701 CCACCCCGCG CCCTCCACTC TAGTTCCCAA TTTTACAATT TCCCATTTCA
34751 GAGGCAACCA TATTCCCAGT TTCTTTTTTG TTTGTTTGTT TGTTTTGAGA
34801 TGTTTAGTGT ATGATTGTCA TGTGGGGTGA GTGTGTGTTT TTTCCTCCTC
34851 TTTTTTCTTT TTTAAGACAA ATTGTAGCAC TCTGTAGGTA CTGTATTGCT
34901 TCATGCTTTT TTCACTTAAA AAAAGTGATA TAAAACTGTC CCCATGATAG
34951 TGATATGCTA TATCATGTGA TAGAGTGATA TATCATGGGG ATAGTTTCAT
35001 ATCACACCAT CACACCTAGA GTTCTGCCTC ATACTTTGTT AAAAGCTATA
35051 CGGGGGCACC ACGATTTACC TATCGAGTTC CCACTGGTTA ACATTTAAAT
35101 TGTTTTCAGT CTTTCCTTCT TAAATAATGC TGCAGTGAGA TATTTTGAAT
35151 ATAAGCTTTT GTGTATGTGT GTGAGGATAT CTGTGAGGTA AATTTCTAGA
35201 CATGAAATTG CTGGGTCCGA AGGACATGTG GGTTTGTATC CTTGATAAGT
35251 GTCAACAAAT CGCAATGGGA CCATTTTGCA CTCTTGCTGA TGATGTATAA
35301 GTGTGCTGAG CAGGCTTGGA ATGTCTCCTG TCTGTTTCGG CAGGTTGTAC
35351 CAGTCAGTTA AGCTGCTGTA CTCCATCGGG ATCTTTTTCA CCTACGCACT
35401 CCAGTTCTAC GTCCCGGCTG AGATCATCAT CCCCTTCTTT GTGTCCCGAG
35451 CGCCCGAGCA CTGTGAGTTA GTGGTGGACC TGTTTGTGCG CACAGTGCTG
35501 GTCTGCCTGA CATGTGAGTA GAAGATGATA ATTGCCTTGC TTGTTTTTCC
35551 CTAAAGGGCA CCCAGTCTGC AGGCTTTCAT GAGAAAAGAC AATGTGTGTT
35601 GTAGTGAAGC TGGCTATGTT TGTGACAGAG AACCTGGCCC ATGGCCTCAC
35651 TTTCAGAGTT GAGGCACCTC CAGATGGGGA AGTGAATTAA TTACATATGT
35701 ACTGTAAAGA ACATGGGAAT GAGGACAGTG GTTTATGTAT AGATAGGGTA
35751 TGAAATGCTG TGGAGGTGGT TATCATTCAG AGTAAAGACA TGCGATTACT
35801 ATCCCATATT AAATAAGGTA AAGGTCTGAA AGCCATTTAA CCCATATCTG
35851 TAATGAGTAT AAGTTACTCT GATGAAGGGT ACTTATTTGC TTTTTCAAAT
35901 AGTTGTTTTT CCACTGTGAC AAGTTGCTCC TTAGATTTCC TTTAGAGGCT
35951 TTATGATAGT ATTCTAGACA TTTTTTAATG TCAGTCTTAC TAAATATGTT
36001 TCAGAAAATT TCTATTGATT AACCTAGGTA TTTGATTGAT CACTTGTGTT
36051 TTATTCTTCT TCTCTCAACC CCATTCCCAG GAGTGTAAGT TAAAAGACAG
36101 GATACCCTTC TGTTTGCTGT GGTTGAAAAC TGGTGACATT TAGAAAATAA
36151 AAGTAAATTT TTTTTGTAGC TTCTGTGAGT TGGTAGACTA GAGAACCCCT
36201 GAGCAAATCG GTTGATAATA GCTAATTTAA GTTTCTAAGA GATTTGCAAT
36251 TGTTTTCCAA ATTCAAATGC TTAAAAGCAT AGATTCCTCT TTTTGGCTCT
36301 ATTTGGCTTT TTTTTCTCTT TTTAGGTTTT ATTATTTTTG AACAAGAACC
36351 TCTTTGCTTA TTATGTTGAG ACTTCCCTGA GAATTTTCTT AAATTATTCA
36401 GTCTGAGCCT CTGTCTTTGG GATAAAGATA GATCCATATG ACTTTTTAAA
36451 TTCTAATTAG GGTTGAATGT TTTAAGGATG AAAGATGGGA AAGTTGTCTA
36501 GCATTTGCTC TTAGTCACTC CTTCAGGCCC TCTCCTAGAC CAGCCTATAT
36551 AGAAACAGCC CACGCAGCAG CTAATCCAGG GGCCAGGGCT GTTGAAAGCC
36601 AGCTGCTGTT CCCACAGCGA CTGAAAAAGA AGGAACATGA TGTATCCTGC
36651 TTTTCTAATA GATTGCCTTA ATGTGTGCTG CTAAGATGGG ATGCTTGGAC
36701 TGTAAATTTT AATCCTATCT TGTGCCAGTA ACTCTCCATG CTTTGATTCC
36751 AAAGTGTATG TTTCCACCGT GGATGGAGTA GCTCTAAGTG CTTGAGGAGA
36801 CAGCTTTCAC GTGTATGGTA TTTATAATGT AAACTCTGAG GGCCCAATTC
36851 TTAAATCTAA AGGGCACTGG AAGAAAGAGT GTGGTTAGTT CAAATAATTT
36901 GCTTTTATCC AAAGTGCTCC CTCCGGAAAA AGTAGGTCTC TGTAGGTAAA
36951 ATGTGCCTTC CTGACTAAAC AGCTCCTCCA CCCTGCCTAT TGAGCTGGGG
37001 CAGTGACAGG AGCCTGACTC CTCTCCCTGC CCAATTTTCC CCTCCAGCCT
37051 GGCTCAGCCT CCCTGTAGCA TATGTCACAC TTCCTGCCAG GTTTATTTCT
37101 GCAGCACCCT GCAGGAGACA GCAGTCTCTG ATTCACAGAC CTCATGTTAT
37151 CCTTAGATGC CTCTTGGATT TTGCTTCACT TTTCCTGGCC CTGTCTGTGA
37201 GTCTCATCTC CCTTCAACAG GACGATGCTC AGAAGACACG GCTGCTTTTG
37251 GTCTTCAAGT GTGTGCAGTT GTTTTTCCCT TCTGTGATCT GTTGTGACTT
37301 AGCATTGCAT TGTCATCCTG TTCAAAAAGG CAGCCCCCTT TATGTCTGAG
37351 AGCACTCGCC TCTCTCACCT TCCTTGGAGA CTTTGAAGTA ATTGTGGGAC

FIGURE 3K

37401 TCAGTAGAGG CCTTTCATGG CAGCAGCAAC TTAAATGTAT TTATGCGCGT
37451 TCATTTTGTT CTTGCTTCTC TGTTCTTTCA GATCTTTCAG CACCGTTGGT
37501 TAGTATGTGA TTTTAGATCT TTAATTGATT TTTTTCATTT ATATTCATAA
37551 ATTTTAACAG CAGCTTCTTT TATTACTATT TCTGGTGTTT TCCTATCTTT
37601 TCCCCAACTT TTCCTCCTCC TCTTCACCCT CCAAAGGGAA CAGGAGGAAT
37651 TCAGTGTAGT TTCTTTTTTT TTTTCCCTCT TGGAATTCAA CTTTCTCACC
37701 ACTCTCCCCC ATCCTCCAAA GATTACTATG GCTGATACGG ACTTTGTGAT
37751 GCTTAATTTC AAACAGTTGG AGAAGAGGGG GAGGGAAAAC AAGTATTTCA
37801 TAGGATAGTG CTCATTTTGT TATGATTTCA TATCGGACAG TATCTACTTC
37851 CAGCCCATAT TTTTGGAAAT GCGGACTTAG CAGGTCACCT TATGTCCAGA
37901 CCTTGTGTGG AAGAGGCTGG CCCCACCTGT GGAGTCTGGA GTTGTAGGAT
37951 CAACGGTTTT TTAGATTTCT TTGGAGCAAT AACCCATCCA TCCTTCAGTG
38001 ATTCATACTG ATTCTCTGTG TCATTTGCCA TGTGAAACAT TTTACTTCAG
38051 TTTGCTATGA AAATTTCAGA AACCTATTTC TGAAGATATA ATTACCTAAA
38101 ATCGCATCAT CCAAGAAGCC TGTTCAGACT GGAATGCAGA GCTGCAAAAC
38151 ATTCCAAGCA GTCGGATTTT TAGAGGATGA AGCTTCCAGG TCCAAACAGA
38201 GTAGCTTCTT AGTACCTTTG GGCCTTTCAC ACTTTTTAGT CTTGCAGCTA
38251 CAGTGAAGAA GAGCAGCATC ATTAATTAGC TGTGTAACCC TGCCACCCCC
38301 CACCCTGCAT TCCCCGCCCA GGAACCCTCA TAAGGCCTCA GGGTCCTCAA
38351 CTGTAAGATA GGAAGGGTGT CTGACCTCTA AGGTTTCTCT CAACTCCAAA
38401 ATTCTGTGAT TCTGTATAGG TGCTTTGCGC TTGATTTTAA GTTTCTACAC
38451 AAATATTACT CTAAAAAAAG AAAGTCAATG TAAAAACATT TGGGAATAAA
38501 AGAAGAAATT CCAGTATTCC ACCAATTTAA CAAAGTAATT TTTTTTTGCA
38551 TTGTATCTTC TGTGTCTTAA TCCTCATGGG TGCCTTGTAA AAATAGTTGC
38601 AATTGTAGTT TACACATAAT TTTGTCTTTC ACATTTTATT TAGTTTTATA
38651 TCACAAATAT TCATATCTTT CACTAATATT TTCATGACCT CGTGGTATTC
38701 CACTGTATTG GTGGATCATA TTAACTAAGG TACTCCTTTC ATGTTGGACA
38751 TGGTGGTTGT TTCCCTTGTT TTCGTATTTT TTAAATTTAT ACCCCCACTA
38801 AGTCAAACTT TGTATACTGT CCAAGACTAC TGTGAATTTT AAAGGCATAT
38851 TTATAGACAT TTAAAAGTAA CATGGTGAAA CCCCGTCTCT ACTAAAAATA
38901 CAAAAACAAA ATTAGCCTGG TGTGGTGGCA GGTGCCTGTA GTCCCAGCTA
38951 CTTGGGAGGC TGAGGCAGGA GAATGGCATG AACCCAGGAG GCAGAGCTTG
39001 CAGTGAGCCA AGATCGCGCC ACTGCACTCC AGCCTAGGTG ACAGGGCGAG
39051 ACTCCATCTC AAAACCAATA AAAATAAAAA TAAAAATAAA TAAATAAAAG
39101 TAACTTGGTA AGTTTTAACA GCTTTGATCA TAATAAAATA GCAGCAAGAG
39151 CTCCCAGCAC AGGAGCCATA AATGGCCAGC GTATTTCGTA AGTTCGCTTT
39201 TGTTCTTTTC AGTGCTTTGC TCTTGTTGTG TATAAGTCAG CTCTTTCTGA
39251 TGCTGGTTCA AAACCACAGG CTCCAGAATC CAGTTCCTTC TGTGAACATG
39301 ACTGTTGGCC TTATGTTGCT TCAGCAGTTT AAAAGCTCAT ATTCTTTGTG
39351 TCTCTTGACT CGAAGGGAAG ATGTTTTGTA ATACTGTTGG AGCCCTCTTG
39401 ACTAATCATG TGGTCGAGCT GAGGTTGTCC TCTGTCCCCC CTTTTGTACA
39451 CGCCACAGCT GAGCTGCTGC TGAGAAGTGT ATAACTGCAT TTGTTATACA
39501 AATGTCTTCC TTTTTGTCTG GGCTGGGGTC TTTGTGTGTG TGGGGGGGGT
39551 GATTAGGGGA GAGTAGGGAG AGGGCTGTTC CTGGCTGGCT GCTTCCTGAG
39601 ATATCTACCT TGTTGAGTGT CTCTTCATAG GCACTTTAAC TCACAGAAGA
39651 CATTTAGTGC CAGAAGGGGT TTTATTTGCC CCACATGTCT GCATAGTCGA
39701 TTGCTGCTTC TGGAGTTAGT TAAAGTCATT TTCCATGGTG GCAAAACAGA
39751 TACCCGTGCT GTTGAACCCT GGGGGCTGCT GATGCTGATT TGGTTTGGAC
39801 ATCCTTCTCT TCTTCCCACT TTGTGTTAGT GGGAGGCTCG CTCTTCTTGC
39851 CCTCTGCAGT GTCACGCTTC ATGTAGGGTT CAGCGGTGGT ATGTGGTTCA
39901 GCTAGGACAG GAAGAAGGAC TTCCCTTTGC AGCCCTGTGG TCCTGGCTTT
39951 AAGAGGAGAG AAATGTTCTT AAAATCTCTA TTAAGGATAT TTTTATTAGG
40001 CATGTTTATC TTATATAGTG GTGAAAACAA GAACAAGTTT TTAGATTACT
40051 TATAAAATAT CATGATGAAG CGGAAGATCT TTGTCCAATC AGAGGAAAAA
40101 TTCTGATCCC AATCTTCTGT TTCTGTTTCC ACTTAACTCC CACCACAGAG
40151 TGGAGCATCT CTCTGACTCC ACTTAACTAT CATGAAGTGC CCATATGTCC
40201 TGCTAGGTCA GTATGGGAGA GGGTGGGGAG ATGACAGACT CTCAGGGCTG
40251 GGAAAGGCTC TGATTTGTCT CCTGCCAGGG ACTCATTTTC TCTGATAATA
40301 AAGGCCTCCT GTCTCTTGAG CGGACATCAG CATTTGTGGA GATGCTTGTG
40351 TGGCTGGGAC TGAAGGAATA ACTCACACTT CCTTTATCCA TACAAAACCG
40401 AGTGGGTTAG AAGCTCCCTT TTGGGCAAGC CATGTGTTCG AGGCTTGGAG
40451 CCCCATCGCT TTGCTGTGCC ACCCTCAGGC AGGACGTGGT GCTTCCCAGT
40501 TGTCAGTGAG GTGAGGAACA TATCCCAGAA CACAGTCCTA AGTGACTAAC
40551 ACTGGAGTGT ATAGTTCCTT AGAATTTCAG AGTTGGGCGA GACTTCAGAC
40601 ATCACCCAGT TACCACATTT CACAGGTAAA CGAATGAACT GAGGCTCAGA
40651 ATAGTAAGCT GATTTGCCCT GCACCACTCA GCTTGTTATG GAGCAGGGAC
40701 TGGTGATAAT ATTGAAGCAT TTATTATTGG TTTTTAGAAC GCTGAGTTCT
40751 TTACATGAGT GATATCGTTT GACCATCCTG TTTAGTAGTT GGGGAAATGA

FIGURE 3L

40801 GTCTCAAAGA GTTTAGGTAA CTAGCCACTG AGTGGTAGAG CTGAGGTTGG
40851 AGCCTGGGCA TTCCAAATCC AGAGCCTACA CCCATTCCTA TACCACCCTC
40901 TCCCTGGGGC TCAGTTCTCC CGATTGTTAC CTCAGTGCAG CTTCCACCCC
40951 ACAGTAACAC CTGGCACCAT CACTGCAAGG ATGACATTAG GCAGGGAGAC
41001 CCAGACCCCA GAGAGGGCAA GTGTCTTGGT GTCGGTACCA CAGCAGACCA
41051 GCTCTTTTGT TGAGCGTTAG ATGACCTATG ATGAGAATGC TGTTTTGTCA
41101 TCAGCCTACA ATTTTACCTC CAATTTTCTC CATATTCAAC CCTCAAGGTT
41151 TGGGGGATGT CACCAACTCT ATTTAGAAGC AGCTAAGGGC TGAAAACATA
41201 GTAGTTTTGA GTTTCAGGGA AATTAAAAGC CTAAACTTTT TGCACACTTC
41251 TCAAAGCCTT GAGATCGGTG AAGATGTTAA TGAGAATTGC CATTATGGTT
41301 GATTAATAGA AGAAGGAAAG ATGAAGAACT GCCCCAGAGT ATATAACACT
41351 GTGGCAGAGG TGGATCTGAG ACTCGATATC CTTGTACCCG TTTGCTGTCA
41401 GGCTTTGGAT CCCTTGTCCA ATTCCACTTG ACAAAGCAGA AAGAGGTCAG
41451 GGCTGATCGT GTGCTGGGTT CTCCACCATG CACCATGGTG CATCCCTGTG
41501 AAAGCTAGCC TGGGTATCTA CCTCTCATTT CCTCACAGGC AGGATCCTTC
41551 CACTGATCCG CCAGACTCCC TGTCTCCCCT CCCTCTGCAT TTCCTTGCCC
41601 ATCAGGTCTG TGAGGTTATG GGCCAGGGGC TTGAGGTCCT GGATCCTGGT
41651 CCCAGCTCTG TTGCTTCCTG TCGTTTACAC CCTCTGGGCC TCTCTTTCCA
41701 TAGGACTGTA GTAATATTGT GGAGTTACAT ACCTGTGAAA CAGAGGCAGA
41751 TTCACCTCCA CTAGTGAGTG CTTAGCAGTG CTCTCTGCTG GGTACCACTA
41801 GACATTCTGC AGTAATGGAA ATGAATGGAA ATGTCCCATG TGTGTGCTGT
41851 CCATTGCAGC AGCCACTAGG CACCAGTGGT TGTTGAGCCT TTGAAATGTG
41901 GCTAGTGTGA ATGAAGAACA GGATTTAATT TCATTTTAAT TCATTTCAAT
41951 GTAAATAGCC GTCCGTGGCT ATGTTGGACA GCACAGCTCC AGGGTAAGTG
42001 TGAGGCAGGA GGCATGAATC CATTCTTTCC CTGGTGTGTT AGTCCATTTG
42051 CGTTGTTATA AAGAAGCACC TGAGACTGGG TAATTTATAA AGAAAAGAGG
42101 TTTATTTTGG CTCATGGCTC TGCAGGCTGT ACAGGAAGTG TGATGCCAGC
42151 ATCTGCTTCT GGTGAGGGCC TCAGGAAGCT TCTAATCATG GCAGAAGGCA
42201 AAGGGGGAGC AGGCTTTATA TGGCAAGACA GGGAGCAAGG AGAAGGGAGG
42251 TACCAGGCTC TTTTAAACAA CAGCTCTCTC AGGGAGGGCC CCAAGTCATT
42301 CATGAGGGAT TTGCCCCCAC GACTCAAACA CTTCCCACCA GGCCCCACCT
42351 CTGACATTGG GGATCACATT TCAACATGAA ATTTGGAGGG GATCCAAACC
42401 ATATTACCTG GTAAGTCCTT GTTTCCACAT GTCTCTCATC TTACTGCAGG
42451 GAGTGCTATT CTCTTTTGTT TGTTTTTATG GCTCCTCAAA AATCAACTTT
42501 AGACATTTCA GTTTAAAGTG TTTCTTAAAA ATCTGGTCTC TAAATGCAAT
42551 CCAATCCTTC AGCTGCTCAG CCAAAGAAGC AGTGATCGAT GTAGACATTG
42601 GCTGCCTTGG ACTGAGATGT TCTGGCAGTC TCACCAGTGT GGTGCCTTCC
42651 TTAGAGTGAC TTGACTGCAT TTTCGCTTTA CAGAATGAAC TTAGAAGCAA
42701 ACCTCTCATA TAAAATGTAA CCCTCTCGTA GGAATCAATG AGGTAGTAGA
42751 TAAGCTCTGG ATGTCTGTAT CAAGGCTGGG AGCATCCAGC TGTAGCCCAG
42801 CAGTAGGAAA GACAATCTGT CAAACTATAT TTGATTGCTA ACAGGTTAGT
42851 AACTAACAGG AAGTCATGCA CTGTAGCAGG ATGTACTTTT CATGGCCAAA
42901 AAGATGAGTA CTAATGATGA TAACATTAAC AGGTAAGACA TCCCTACTGT
42951 ACACCAGGCC TTTTGTGAGG CACCTGCATA ACCTCATTTG ACCATCATGA
43001 CATCTCTATG ATTCAGGAGC AGTTAATATC CCCATTTTGC CAACAAGAAA
43051 ACTGGGGAAT AGAAAGGTAC CATACCTTCC CCAATGTCAC TCAGCTAATT
43101 AGCAGCAGAG CCAGGATCTG AACACAAGAA CCTAGTTCCA GAGCCCACAG
43151 GCCTCAATAA ACCTGTGAAA CACTGGCCTT TGCCCACCTG GTGGAAAGAT
43201 CGGTGAGATG GGAAGCGTGG GGTCAGTGGG CACTAGGATG GGTGTATTCG
43251 GTGAAGCCTC CTCCTGCTTA CAGCACTGTC TGGCAGTGTT GACAATGGCT
43301 GGTATGGCAC GGAAGCCGAT GGCACCTCCT GCGGCAGTGC ACCATTGGTC
43351 TTCGTCAGTT CCTCCTTCCT GGCTCACCCG TGGCTGAGTT TCAGATGTGA
43401 GAGCCAGTGG GTGTCCTGTC ACAGAGATAC GGTCGTCGTG TGGGGCTTCG
43451 CCAGGGGTCA GCCTGCAGAT AGAACTGCTT TTTTTCACCT GTATCAAAAT
43501 GCTCTGTGAA ATGCGGTTTT ATCACGGTGT CTTTCCAGAA GGCGGGGTTT
43551 CTTTTCCTAT TTGGTTTCTT GTCAGTCAGG TAGAGATGTT TGTGTTGGAG
43601 GCTCCCTGAG TGGTAAGAAA ATGAGCAGCT GCTCAGGAAC GTCCACCTCC
43651 TTTTCTTCTC CCTACCCTCC CTCCTTGGGT AGAAACCAGA TTGACCTGAA
43701 ATGTAATTTG GTTCCTTTTG CACAGAAAGA TGGAAGTCAT GTGCTGTAGC
43751 CGGAAAAGCT GAAAGCCTGG GGACCGGAGC CAGAAGATCC GGGTTCCGGT
43801 CCCAGTTCTG CTGACCTTGC AGTGGGACGC CAGTTATTCC ATGTTTCTGG
43851 GCATCTATTT CACAGAGATT GAACTGGACA ATGTCTAAGT TTTCTTAGAG
43901 CTCTCAATCC TATAAGATGG ACAGATGCTG AATTCGCTCA ACAGTAGGGA
43951 GACAGACCTT TCCCAGATTC TGGGCATCTT AAACAGGCAT GTCCTCTCTC
44001 CCTGCAGGCA TCTTGGCCAT CCTCATCCCC CGCCTGGACC TGGTCATCTC
44051 CCTGGTGGGC TCCGTGAGCA GCAGCGCCCT GGCCCTCATC ATCCCACCGC
44101 TCCTGGAGGT CACCACCTTC TACTCAGAGG GCATGAGCCC CCCTCACCAT
44151 CTTTAAGGAC GCCCTGATCA GCATCCTGGG CTTCGTGGGC TTTGTGGTGG

FIGURE 3M

```
44201 GGACCTATGA GGCTCTCTAT GAGCTGATCC AGCCAAGCAA TGCTCCCATC
44251 TTCATCAATT CCACCTGTGC CTTCATATAG GGATCTGGGT TCGTCTCTGC
44301 AGCTGCCTAC CCCTGCCCCA TGTGTCCCCC GTTACCTGTC CTCAGAGCCT
44351 CAGGTATGGT CCAGGCTCTG AGGAAAGTCA GGGTTGCTGT GTGGGAACCC
44401 CTCTGCCTGG CACCTGGATA CCCTGGGCCA GGTAACCTGA GGGCAGGGGA
44451 GAGGTGGGGT GGCAGACACG CAGAAGTGCT ACTAGTGACA GGGCTGCCAT
44501 CGCTCACCTG TACCTATTTA CACCCAGAAC TTTCCAGCTC CCCCTCATCA
44551 TGCCTCCTCC TTCCTACCTG CCTCCCCTCT GCTGGTGCAC CTCGCCCAAC
44601 TCATTCTTAC TGCACAGTTC ACTTTATTTA ACAATTTTCA TGTCCCCCAC
44651 CTCATGTTTT CACCTTTTAC TGGGCCAGGC ATAGATTAAG TAACTGGGAA
44701 CGCCCCCTCT TTATAAAGCT GGGCTTCTTT CTCATCTCTC TCCCAAATGT
44751 TGTATACTCA GTATTCTTCC TATTCGAGTC TCCAGGGGGT GGCTGGACCT
44801 ACCTGGTCAT TTGAAACAGG CCCCCAAGCT GGAGTTTTTA ATCTGGACTC
44851 TCTGGCTTGC TGTGACCCCT AAGGCAATGC TTCTCTTCCC TGGTATTCCT
44901 TAGTGTGGGT CACAGTACTG TGTTCTTAGT TGCTTTAGCT CTTAAAACAT
44951 ACGAAGTGTT GCCTAAACTG AAAATATTTA TCTTTTATTT AAAATCAGAT
45001 TTTTGTTTTT AGACTGTCTT AGATCTGGGG CTATTACGAA TCACTTCTTC
45051 TTCAGTAAAC TTTGACTCAA CTTCTCCTGC TGAAAAGAAG CTCGCTCCAG
45101 ATGTCTGCAT GGGTCCTCGG CACTCTTGGC TGAGGACTCA AAGGTTTTAA
45151 TCAGGATCGT CTAAAAATGT ACCTCGGTGA GGAGGCACAG ATTTTGCCTC
45201 CTGTTGACCA GCCTGGTTTC ATACCGAAAA GACATTGAAG GACTGCAGAA
45251 ATGTATGGGT GCACCGGGCC GAGGGAAGGG TGGCTGAGTG AGAGGCGTAT
45301 AAAATGGGGC TGTGTGCATG CAGGCCCATG TTTCAGCCTC AGCCCACGCC
45351 AGGTGAAAGG ATCAGCAATG CTCTGTTGCC ATCGTGCTGG GACGACACCA
45401 GCTCTATTGC CACCGATGAG TAGCTGAGGT CAGTGTGCAC AGAGTTTGAA
45451 ATTAAGTTAA TAGACTTTAC AGCAGCTGGT CTGACACTAC GCGCAGTGCT
45501 CGGTTGTTTA CAATCAGTGG GGAAAAGGGC AGAACCAGTG CCCGGCCCCA
45551 CACTGCCTCT GTGGCCTGGA CTTTGAAAGG AACCCACTGA ACACTAATTA
45601 TGAGCCCTGT CTTTCCCCCA GAATGCCTCC CTGGGTTTCA CAAACAGCCT
45651 TGAGGTTGGC CCTCCTCAAG GTCAGCCTTC AGATTTGGGA GCAAACTTCA
45701 GAGAAGGCAG AGGAAGATAC ATTGCCTTGC TGTGGGCTGC CTCTTCTTTC
45751 CTCTTGGTGT GCGAAGTATT TCAGAAGGCC ATTGATGAAT TCCCCCTCTT
45801 TAGCTGTGTA TTTGTGCACG TGTGTGTGTA CGTGCGTGTG TGTGTGTGTT
45851 CCTGTGTAAG TAACAGACCA GACTCCTTTT CTCTTCTGTC CCGTCACCAG
45901 GCTCTTGCTT CACTGCAGAT ACAGTTCACT CTGAAAGCTG GTTGAAGGAG
45951 AGCAGCAAAA ATGTATCAGG GGTTTTGCTT CTGTGTTTCG CCAAAGCTCA
46001 TAAGGGCTGT GACCCACCCA TATGGCCCCA GTTTTTTCTG TCTCTTCTGT
46051 TCCAAAGCCA GGAGAGCTGA CTTCCAGGTG AAGGGATGGG AAAAGTGGAC
46101 TCTCATTGTA GTGACTCCCA ACCTACCTAA TAATTTGTTA ACTTAGGAAT
46151 ATGCTATCAT TGTTGACTTG TTCTTCCTTA GGAGAAGGAC GATTTTCACC
46201 CACCCTTTCT GTTCTATGGT GGACTCTTAA CAGGTGCTAT GTGACCAGGA
46251 ATCTAGCCGG GAGTAGCAGA GGCCCTGTCT TCTGAAGTCT CAGGCTTAGA
46301 AGTTACCAAA GTGGGCTCAG AAACTGTCAT CTCCTGGTTC CAAGTTCGGG
46351 CTCTGGCAGC CCAGCCGCTA TCTTAGCTGT CTTTCCCAGC GGTGCTAAGA
46401 GTGGTCTCAG TGAGAAGGTA GATGCCAACT GGAGGGCCAG ACCTGTGTCC
46451 TGCCCATGTC CTCCTTGGTG GACGTTTCTG TTTACTCAGA GCTGCTAGAG
46501 ACCATCCTGC CCATCCGAGT TCTGAGATTG GGACTGTGAT GTTGGGACCT
46551 GAGGACTGGA TGGTAGAATA CTGGGGTCCC CCAGCTCTTA GCAGGATGCA
46601 GGCTATTGCT TCCACACCCC TGGCCGTGAG AACGTGGTAT GTAGGAGAG
```

(SEQ ID NO:3)

FEATURES:

Start:	13181
Exon:	13181-13323
Intron:	13324-17943
Exon:	17944-18034
Intron:	18035-20533
Exon:	20534-20622
Intron:	20623-21093
Exon:	21094-21189
Intron:	21190-23220
Exon:	23221-23305
Intron:	23306-23728
Exon:	23729-23947
Intron:	23948-29696
Exon:	29697-29795
Intron:	29796-32613
Exon:	32614-32780

FIGURE 3N

Intron: 32781-35343
Exon: 35344-35513
Intron: 35514-44007
Exon: 44008-44277
Stop: 44278

CHROMOSOME MAP POSITION:
Chromosome 5

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain
2064	G	A	Beyond ORF(5')
2119	A	G	Beyond ORF(5')
2121	G	T	Beyond ORF(5')
2123	G	C	Beyond ORF(5')
2125	T	G	Beyond ORF(5')
2825	G	T	Beyond ORF(5')
3288	C	T	Beyond ORF(5')
6172	A	G	Beyond ORF(5')
6462	G	T	Beyond ORF(5')
7031	C	T	Beyond ORF(5')
7671	T	-	Beyond ORF(5')
8466	-	T	Beyond ORF(5')
9097	-	T	Beyond ORF(5')
9108	-	T	Beyond ORF(5')
10170	A	G	Beyond ORF(5')
10966	A	G	Beyond ORF(5')
12987	A	C T	Beyond ORF(5')
13111	C	G	Beyond ORF(5')
13120	C	T	Beyond ORF(5')
13822	C	G	Intron
14891	G	T	Intron
15207	C	T	Intron
16162	G	T	Intron
16364	T	-	Intron
16411	T	C	Intron
16636	T	C	Intron
16802	C	T	Intron
17111	A	G	Intron
17276	T	G	Intron
17372	C	T	Intron
18317	T	G	Intron
18342	C	T	Intron
21828	C	T	Intron
22674	T	C	Intron
22683	T	C	Intron
22822	C	G	Intron
23073	G	A	Intron
23343	C	-	Intron
23396	T	A	Intron
23511	G	C	Intron
23522	C	G	Intron
23582	T	A	Intron
24977	G	A	Intron
25131	A	-	Intron
25178	G	A	Intron
25351	G	A	Intron
25380	A	G	Intron
26737	G	A	Intron
26829	G	A	Intron
27423	G	A	Intron
27735	G	A	Intron
29875	C	G	Intron
30356	G	T	Intron
31344	C	T	Intron

FIGURE 30

32570	T	G	Intron
33220	T	C	Intron
33525	T	G	Intron
34589	A	G	Intron
34832	G	T	Intron
35188	A	G	Intron
35614	G	C	Intron
37852	C	A	Intron
38643	G	A	Intron
39198	G	T	Intron
39550	T	G	Intron
42281	A	G	Intron
42321	G	A	Intron
42563	G	C	Intron
42675	G	A	Intron
42908	G	A	Intron
43358	-	G	Intron
43371	G	C	Intron
44796	G	A	Beyond ORF(3')
45820	A	G	Beyond ORF(3')

Context:

DNA
Position
2064

CTTCGACTGGACTCTGCCCCATGCCCAAGATCAATGCCCTGTTCAGTTCCTATTCGCAGT
CCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGTGGTGC
TGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGGCGCGGCCGCTCAAGGGAGCACGT
GACCTCGGCCTCTGGCGTGGGCGGTGGGATCACGTGATGAGGTCCGGAAGCGGCTGCCGG
GCAGCAAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTGCTGAA
[G,A]
CCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCACTGACGAGT
TCCGGCTGGCGGCGCTCGCCGCCTTGGGCAGGACCCACCTCGCCTTCCTCCCGGCGTGGC
AGATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCGCGACTCCTTGGCGTC
CCCGGGCCGCAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTGGAG
ATCCCTTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGGTTGGGATCTCACCGTCCTGT

2119

GCAGTCCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGT
GGTGCTGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGGCGCGGCCGCTCAAGGGAG
CACGTGACCTCGGCCTCTGGCGTGGGCGGTGGGATCACGTGATGAGGTCCGGAAGCGGCT
GCCGGGCAGCAAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTG
CTGAAGCCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCACTG
[A,G]
CGAGTTCCGGCTGGCGGCGCTCGCCGCCTTGGGCAGGACCCACCTCGCCTTCCTCCCGGC
GTGGCAGATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCGCGACTCCTTG
GCGTCCCCGGGCCGCAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCG
TGGAGATCCCTTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGGTTGGGATCTCACCGT
CCTGTTTGACTGACAGGTGGGGGAAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCA

2121

AGTCCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGTGG
TGCTGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGGCGCGGCCGCTCAAGGGAGCA
CGTGACCTCGGCCTCTGGCGTGGGCGGTGGGATCACGTGATGAGGTCCGGAAGCGGCTGC
CGGGCAGCAAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTGCT
GAAGCCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCACTGAC
[G,T]
AGTTCCGGCTGGCGGCGCTCGCCGCCTTGGGCAGGACCCACCTCGCCTTCCTCCCGGCGT
GGCAGATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCGCGACTCCTTGGC
GTCCCCGGGCCGCAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTG
GAGATCCCTTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGGTTGGGATCTCACCGTCC
TGTTTGACTGACAGGTGGGGGAAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCATC

2123

TCCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGTGGTG
CTGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGGCGCGGCCGCTCAAGGGAGCACG
TGACCTCGGCCTCTGGCGTGGGCGGTGGGATCACGTGATGAGGTCCGGAAGCGGCTGCCG
GGCAGCAAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTGCTGA
AGCCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCACTGACGA
[G,C]
TTCCGGCTGGCGGCGCTCGCCGCCTTGGGCAGGACCCACCTCGCCTTCCTCCCGGCGTGG

FIGURE 3P

```
        CAGATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCGCGACTCCTTGGCGT
        CCCCGGGCCGCAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTGGA
        GATCCCTTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGGTTGGGATCTCACCGTCCTG
        TTTGACTGACAGGTGGGGGAAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCATCTA

2125    CCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGTGGTGCT
        GTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGGCGCGGCCGCTCAAGGGAGCACGTG
        ACCTCGGCCTCTGGCGTGGGCGGTGGGATCACGTGATGAGGTCCGGAAGCGGCTGCCGGG
        CAGCAAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTGCTGAAG
        CCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCACTGACGAGT
        [T,G]
        CCGGCTGGCGGCGCTCGCCGCCTTGGGCAGGACCCACCTCGCCTTCCTCCCGGCGTGGCA
        GATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCGCGACTCCTTGGCGTCC
        CCGGGCCGCAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTGGAGA
        TCCCTTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGGTTGGGATCTCACCGTCCTGTT
        TGACTGACAGGTGGGGGAAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCATCTAGG

2825    AGAACTGCAAGGACGCTGGGAAGTCGTCTGGTGCAGCTCCCTCCTAGGACAGTTGGAGAA
        ACTGAGCCCTTACTCCGGGAAGGGGTAAGGGCTTGCCTAAGGTCATCCAGTGAGTTAATC
        GGAGACCCGGAGACCTGCGACTAGAATGCAAATGTTCCTAAGCTTCAGCAGCTGTTTGCT
        TTTCGCCACACCGCCTCCTGCGGGAAACTTCACCTGTGAAAAGGCACTCCTTTCTGTCCC
        TTTCTCTTTTAGTCCTCTCCCTTTTTAGCTGTCTGCATTTTCCACCGCTGGGGTTGGATT
        [G,T]
        GCTCTGGGTGTGGTTCCCTGTTTGTTCATTATTTTTCTGCAAACTCATCCTTCTGTAGGT
        TTGGTTTCTAACCTTCCTGCATTCTATGTAAGTCACACCAAAATATGAAATATGAATCGG
        AATGTGCTTCTGGGAAGATAGGTGGCTGAGCCGAGGTTGTGGAGAGCCCTGACGTTAACT
        TGAAGAATGTAAAGACCTTTGCTTTATTTTTTCTGTAACTTGTCAGATTTGGGATTGCTT
        ATTTGGATGGACGTTTTGCAGTTATTTGAATTTTGCTGAAGATAGCATCATGGTGCAATG

3288    AGAGCCCTGACGTTAACTTGAAGAATGTAAAGACCTTTGCTTTATTTTTTCTGTAACTTG
        TCAGATTTGGGATTGCTTATTTGGATGGACGTTTTGCAGTTATTTGAATTTTGCTGAAGA
        TAGCATCATGGTGCAATGGACAGAACAGAGATTGGGGAATCAGGATATTTTGTCCTAGCT
        CTGCCGCTTACCTGGCAACCTTAAGTGACTCGCGTTTGGGTTTCTCAGTCTAGACAGTGA
        TGGAATTGAATTCTTAAGGGCCCCTTCTGCTGTGATCTGGATGTTGTGCATCTTTCTAGG
        [C,T]
        TTGTTTTTTTGTTTGTTTGTTTTTAAATAGAGATGAGGTCTCACTATGCTGCCCAGGCTG
        ATCTCAAACTCCTGGGCTCAAGTGATCCTCCCACCTTGGCCTCCCAAAGTATTGCGATTA
        CAGGGGTGTGAGCCAGTGCCCCTGACCAGGGTCTGTTTGTTTTTTTATTCCGAGAGATTT
        TACCCGCTGTGTACACTGAGTATCAGCCTTGCAACAAGACTTAATCTAATTGTGTAGGAA
        GCAGTTTCCTCTGCTTATTCCTCTGTTGCTATAAAATCCTCCTCCTCTTTCTTCCTATCT

6172    GCTTCACAATCCAGAGTTTTAAATGGAGCCCATACTGCAGACTTTGTTTAGTGAACTTTC
        TCACTTTCTGTCCTTGAACTTCTCTGTAGTAATAATCACAATTGCTATCATTAATGAGGG
        TTTATTATGTTCCAGGCAACATATCTAACATTTATTTATTTTTTCCTCCTATCTTCATAA
        CAATATTGTGATGTAGATGTTATTAATGACATCTTTCAGATGAGGAGACTGTGGCAAAGG
        GAGATGAATTAACTTGCTCAGAGTCACACTACCAGACTGCAAACTCAGGTGCTTTTTATT
        [A,G]
        TGCAGAATACCCGCTGCAGACCTAATCCTGCCCCAGGCTCTGGGGCCAGCTTTGTTCGCA
        GGGAGATTTTAAGGAGGGTATATAATTTAAGGTGTGGTAGAAAGAATACTGGACTGGGAT
        GCTGGTTTGCTGGACTGTCATCTCAAATCTTTGATTTGACTCATCCTGGGGCCTGGATGA
        GTCAGCCTTTGTGTGTGGGCCCTGGTTTTCTGACCCCTAAGAAGGAAGCTGGAGCTTGAC
        CTTCTCTAAAGCTATACCTGGCCCTAACATTTAGTGATCTTCATGGTTGGGAGTAAAAGT

6462    GCTTTTTATTGTGCAGAATACCCGCTGCAGACCTAATCCTGCCCCAGGCTCTGGGGCCAG
        CTTTGTTCGCAGGGAGATTTTAAGGAGGGTATATAATTTAAGGTGTGGTAGAAAGAATAC
        TGGACTGGGATGCTGGTTTGCTGGACTGTCATCTCAAATCTTTGATTTGACTCATCCTGG
        GGCCTGGATGAGTCAGCCTTTGTGTGTGGGCCCTGGTTTTCTGACCCCTAAGAAGGAAGC
        TGGAGCTTGACCTTCTCTAAAGCTATACCTGGCCCTAACATTTAGTGATCTTCATGGTTG
        [G,T]
        GAGTAAAAGTGTGCGTGTTTGCCTGTTCAGCAGCTGCTTTGTGCAGAACCTGCTGAGGTC
        AGCAGCTGCCCTGTAGCTGTTCTAGCATCAGACTCCTACAGGAAAAAGTCTCAATTTATG
        GAATGTTCTGCTCTGGTAAGTTGGATGGAATTCTATCTGATGCTGTTTTAAAAACAAATT
        ATGTAGAAGCCAAACCATTTTACTTCCCTCACTGTAGACCACACATAGCAACACAGTCTG
        TGTCTTTGTTCATGTTTTTAGAATTCCATCGACAGAGAGGAGAAAATACATCTGGGGAAT

7031    TCGACAGAGAGGAGAAAATACATCTGGGGAATTTGCCGCTGCTCTGAGTTCCAAAGTCCA
        AACCAATGTAATTGTTTCAGAATAACGGATGACACTTTTAGCTTGCAAACAAGGGGCGCC
        AATGCGTGAATTCTGGTAGGAGGTGAGGCCTAGGGTGTACCTATCATAATAAGATCATAT
```

FIGURE 3Q

ATTTTTTGTAGTGCTTTATATAAATCTACCTATAATCAAGATTACCTAGGAAGCTAGTTA
AAAATAAAACGCCTCTTGCCTGTAATCCCATCACTTTGAGAGGCTGAGACAGGTGGATCC
[C,T]
TTGAGGTCAAGAGTTTGAGACCAGCCTGGCCAACACGGAGAAACTCCATCTCTACTAAAA
ACACAAAAAATTATCTGGGCATGGTGATGGACGCCTGTAATCCCAGCCACTCGGGAGGCT
GAGGCAGGAGAATCGCTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCCAAGATCACACCA
TTGCACTCCAGCCTGGGCAACAGAGGGAGACACCATCTCAAAAAAAAAAAAAGAAGACAA
AAAGACAAAAACAACAACAAAAAAACATAGGCTGGGCATGGTGACTCATGCCTGTAATCC

7671 TCAAGAGTTTGACACCAGCCTGGCCAACATGATGAAACCCCGTCTCTACTAAAAATAGAA
AAAAATTAGCCAGTTGTGGTGGCGCATGTCCGTAATCCCAGCTACTCGGGAGGCTAAGAC
AGGAGAATTGCTTGAACCTGGGAGGCGGAGGTTGCAGCGAGCCAAGATCGCACCACTGCA
CTCCAGCCTGGGCAACAAGAATGAAACTCCATCTCCAGTAAATAAATTAAAATAAATAAA
TAAAATAAAATAAAATAAAATGCTAAGGTGGAATCAAGTTGGGCCCAGAAATCTATTTTT
[T,-]
TTTTCCTTGACGTATGTTTCATTTAACCCAATATATCCCAGATATTATCATTGCAATATA
TAATCAGTATAAAGATTATTAATTCATGGGATATTTCACAATTTTTTTGTTACCAGTTCA
TTGAAATCTAGTGTGCACATTTCAATTTTACCCAAGTGTATTTCAAGTGTAAGATAGCTA
TTTATGGCTAGTGGTTACTGTACTGGATGGTACAACTTCAGAATATGTTACCATCTATTG
ATCTTAATCCTCCTTTATTTTGAACAAACCCAGTCACTAAAAAATTGAAATTGGAATCCT

8466 AAAAAAATCACTGTAGAATTCCCCTTAAAATTGCCCACTTCTGAAAAATTTAACACCTAC
AAATTTTTATTTTTAAAAATAGAATAAAATTTATTTTATTTTTAAAAATAAAAATTCAGT
TTGCACATACATTTTCCATATTGCATCCGTTGCACAAAGTGATTCCACCTGCTCATTTTT
AGTGCCCATCTAAAAATGGCATATTTTGTAGATTGAAGAGCAACACTTGTCTATTTATAC
AGCTAAAACAATAGTTACATAAGGAAAAAAAAGGAATGTTTTAAGGTTTGTACACTTAAA
[-,T]
TTTTTTTTTTTTTTTTTTTTTTTTGGCCATCAAACTTGCAGACTTTTTTTACTCAGTTGCT
CACTCTTCTGAGTCTAAATATCTAATGGAGATTTGGACTTTGTGTTCTGTTTATTGTCCT
CAGTAATCTGAAGGACAAGCTTGCCTTCAACTCTCACATAGTACAACCCTCATTTAGACA
GTTAACAGGTACTATTAAAATCTCCCATAGGGCGGGAACTGGCAATTGCAGCAATAGACT
TGGCTATCAGATTTCATCAAAGGGAGCCTAAGGGCAGTGTGGCCATGGATGCCAGCACTC

9097 GCCTTGGTTTCCAAAAAGAGCCATAGAAAGAACTCCGGGGAGTGGCTCTGCCCACTGTCT
GATGCTTGAATCCTTACATAACTGCTCTGAGAAAGGGCTTTTGCTTGGATTTTTTCAGGG
ATAAGGGAACAGGCTTTCTCCCAGAGTGATCTGTTCTATTTGGAACAGATCTGTCTTTGA
TAGAAAGTTCTTCCTTACACCTAGCAAAAAATCAGCCCTCTTGACTCTCCACGTACTGAT
CCTAGCCCTGCCTGACCTTTGAGGCCCCAAATAACAAGTCTAATCCATGTGACAGCTTTT
[-,T]
TTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGC
GATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCACACCATTCTCCTGCCTCAGCCTC
CCGAGTAGCTGGGACTACAGGCGCCCGCTACCACGCCCGGCTAATTTTTTGTATTTTTAG
TAGAGACGGCTTTTAAAGACAGTTTTTGTACCCCTCAAGTTGCTAGGTGGAACCTTCTCA
GTGCTTTCAACCATTCCTCATTTAGTTGGTTTCCTACCCCTCTTGATCCTAGTTCTGACC

9108 CAAAAAGAGCCATAGAAAGAACTCCGGGGAGTGGCTCTGCCCACTGTCTGATGCTTGAAT
CCTTACATAACTGCTCTGAGAAAGGGCTTTTGCTTGGATTTTTTCAGGGATAAGGGAACA
GGCTTTCTCCCAGAGTGATCTGTTCTATTTGGAACAGATCTGTCTTTGATAGAAAGTTCT
TCCTTACACCTAGCAAAAAATCAGCCCTCTTGACTCTCCACGTACTGATCCTAGCCCTGC
CTGACCTTTGAGGCCCCAAATAACAAGTCTAATCCATGTGACAGCTTTTTTTTTTTTTT
[-,T]
TTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTC
ACTGCAAGCTCCGCCTCCCGGGTTCACACCATTCTCCTGCCTCAGCCTCCCGAGTAGCTG
GGACTACAGGCGCCCGCTACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGCT
TTTAAAGACAGTTTTTGTACCCCTCAAGTTGCTAGGTGGAACCTTCTCAGTGCTTTCAAC
CATTCCTCATTTAGTTGGTTTCCTACCCCTCTTGATCCTAGTTCTGACCCCTGGATATAC

10170 TATACACAACATAAAGAAGTCTCTCTGCAGTGTTTGAGATAAATTGAACATCTGTACCAA
GTAGACAACAGAGAGGTTTCTCGGTTGCTAGGGAAGGATTGGGCAATTAATAAGTCCCTG
TATTCCATCCTTTCACCTTCAGTAATATATAGGTGTCAACCTAAAGGAAGAAGTTGAGAC
ACAAAATGCAATTTTTAACAGTTTACTTGAACTGTTTACTTGAACCAAGTGAGGACAGCT
GCCCGGGACACACTTCCAAGTTGCCTGGGGGAGTGCGTCCTTCGGCCTTTGTTACCACAG
[A,G]
TTCTTAAAGGCAAAAGCGAACAAGGAGAGGACTGATACAAAGTGACTTGACAGGAATTCT
CATCAGTTTACAGAAATAGCATGGATTATTGATGGGCTGTACATTGTTGGACTATAGGGT
ATGAGTTATGATGTCCAGTGTTAGCATTTTATGACTTAGTGGTGTCAGTTAGTCTAGAAC
CCACATAGCAAGTGGCTTCAAGAGGTAATTATTTAACTCAAGGGGGGAGTGACACATGAC
TGCTCTCACATTTTAGTGCCTCTCTGGACCCGTAATTTAAAGGGATTCCTCAGATAAAAA

FIGURE 3R

10966 CAGGTGTGAGACACTGCACTCAGCTGCTGTTTGCATAAATAATTATGTTCATTGACACCT
AGAATATTAGTGCTAGAGGGAGTTGAGAGATATTTTAGTTTATGCCCCATGCTTTTTGCA
CATTTGAAAATGGTTCACAGGTACTAAGCAAACTGTTGACAGAGGTAGGCTTGGCGCCTG
GGCCTCCTGACATACCTGTAAACTGATTTACGAGCTTATACCTGTATAGCAAGAGGTTAC
AATGCTGGTATTAAGATACTTCAGAGATTTTTTTTTTCTCCCGGCCCTCTAGTGAGTTTA
[A,G]
TTGCCCCAGAGCTGGTTGGCGTCCTTGAATTCCTCTAGCTCATGAGTAAATGAAGCTCTC
ATAGATTTTTAGCCAAGTGGCTCTGGCAATGAAGCTAGGCAGGATCGTCTCTGGGATTTC
CAGGTCCTTTGCTGGCATTTTGCCAGGTACTTCCCTTGTGAGATAGCTTGGGGGTCCTTC
CTACATTGCAATTGTTGAGAGAAAATGCGATCTCCCGTGGATCTCTCTGGTGCCAGACTG
GGGTGTTTCCAAAGGAGTACCCTGGCACTGGACCTAAGGAGAGCCTTCGGCGGAGCACCA

12987 GTGTATATACGTATATATATATATATATATATATGGGATGATTTATATATATATATATACAG
CAGCACGATTGAACTATTGCACAAGGTCCAAGACATTATCTCAGAAAGGAGTAGATAATC
CTGACCTAAGGAATAGGGAATGCGGAATTCCAGGAAGCACTTCTCTTTCATTTTCCCCCA
CTCCTCCCAAGCAGTGCCTCACTTCTGCCTTGTCTAGCTGTACTCCGGAAAATTAAGAAA
TTTATGAGTGTAGCACCACGTATACCAATGGGAAGGATGGGAGTCAGAAGTCAAGTGAAC
[A,C,T]
CAGCCCGCCTCTGTGTACTTTGCACTTTTCCATTTCCCTTGGTACCAGGCACTTTCATAC
TTAATCCATAGTGGAGCTGTCACAGTGAGCAACTCTGACAATGACAGCTTCTACCCCAGA
GGCCACCCCAAACATGGAGCTAAAGGCTCCAGCTGCAGGAGGTCTTAATGCTGGCCCTGT
CCCCCCAGCTGCCATGTCCACGCAGAGACTTCGGAATGAAGACTACCACGACTACAGCTC
CACGGACGTGAGCCCTGAGGAGAGCCCGTCGGAAGGCCTCAACAACCTCTCCTCCCCGGG

13111 CCTAAGGAATAGGGAATGCGGAATTCCAGGAAGCACTTCTCTTTCATTTTCCCCCACTCC
TCCCAAGCAGTGCCTCACTTCTGCCTTGTCTAGCTGTACTCCGGAAAATTAAGAAATTTA
TGAGTGTAGCACCACGTATACCAATGGGAAGGATGGGAGTCAGAAGTCAAGTGAACTCAG
CCCGCCTCTGTGTACTTTGCACTTTTCCATTTCCCTTGGTACCAGGCACTTTCATACTTA
ATCCATAGTGGAGCTGTCACAGTGAGCAACTCTGACAATGACAGCTTCTACCCCAGAGGC
[C,G]
ACCCCAAACATGGAGCTAAAGGCTCCAGCTGCAGGAGGTCTTAATGCTGGCCCTGTCCCC
CCAGCTGCCATGTCCACGCAGAGACTTCGGAATGAAGACTACCACGACTACAGCTCCACG
GACGTGAGCCCTGAGGAGAGCCCGTCGGAAGGCCTCAACAACCTCTCCTCCCCGGGCTCC
TACCAGCGCTTTGGTCAAAGCAATAGCACAACGTGAGTAGCTGTTACCTTCTCCTCTCCT
GGGTGGGATTCGTGTTCCTAAGCCTCCCTTGGACTTATTTTTCCCCCCAATTTCATCAGT

13120 TAGGGAATGCGGAATTCCAGGAAGCACTTCTCTTTCATTTTCCCCCACTCCTCCCAAGCA
GTGCCTCACTTCTGCCTTGTCTAGCTGTACTCCGGAAAATTAAGAAATTTATGAGTGTAG
CACCACGTATACCAATGGGAAGGATGGGAGTCAGAAGTCAAGTGAACTCAGCCCGCCTCT
GTGTACTTTGCACTTTTCCATTTCCCTTGGTACCAGGCACTTTCATACTTAATCCATAGT
GGAGCTGTCACAGTGAGCAACTCTGACAATGACAGCTTCTACCCCAGAGGCCACCCCAAA
[C,T]
ATGGAGCTAAAGGCTCCAGCTGCAGGAGGTCTTAATGCTGGCCCTGTCCCCCCAGCTGCC
ATGTCCACGCAGAGACTTCGGAATGAAGACTACCACGACTACAGCTCCACGGACGTGAGC
CCTGAGGAGAGCCCGTCGGAAGGCCTCAACAACCTCTCCTCCCCGGGCTCCTACCAGCGC
TTTGGTCAAAGCAATAGCACAACGTGAGTAGCTGTTACCTTCTCCTCTCCTGGGTGGGAT
TCGTGTTCCTAAGCCTCCCTTGGACTTATTTTTCCCCCCAATTTCATCAGTCCTCCACTT

13822 TTCAGCTGTGTTGGCCCCCCAAGGCTGGTGCCAAGTGAGAGCTTGGACTTAAAAAAAGCT
TCTACAGAGGACATTCTTTTAATTTAAAAGTGTGTCATCTGTGCTAGAACCCCAAATAAT
TTCCAAGCATAATCGGAAGCTTCCTTTGCAAAGTCTCCCCCCGAATTCTGCCCCATCACC
AAATCAGTATTCATTTGACTGAAGAAGTGGGAAGAGAGAAGAATTAACTTCTGCACTTAA
AAAATTCAGGGTTGGTAGGAAAGGAAAGATAGACTTTGCATTCTCCAAAGAGGGCTTAAT
[C,G]
TCTTGTCTCCAGAAACTGGGACCCCAGACTCATTTGGGCTGAGTTTGGCCCGCTTCAGGT
CTCACTTTCCCCAAATGTAAAGAAAAATTGAGGACTCCACCACAAAGCTATGCTGGCTGT
GTGGGGCTCACCACTTGAATTAGAAAATTCAGAGGAAGTTTTGCTACTCCATTGAGTTAG
TTTCCCAGCTACTCCTGATTTCAGCAGACCTCTGACTTTTCTCTGTGTCCCAGCATCTCA
GCTTTTGCAGTCCTGTTTATTCCTCAAGCTTAGCTATTACCTTTTCTGTGTTTTCTTGTG

14891 GAGGTCATGCTTGTTATCGTAAGATACTTCTAAGCTTCCTTCTCTGTAGTTTCCTTTGCA
GTTTTTGCTCCTTTTTGATCTCAGATATCAACTTGTCTAAGCAATATTTAGCAGATGAGG
TCTGGATTTTTATGTTTATAGAGACATCTCTGAAGCTCAAAACCTACCAACTAGCAACTT
TAGGATAGTAGCTCATAGGTTTTGGACAAAATTATGTCCTTGTTTCTTGGAAATCGAACA
AATCAGAAGATACCTTCCTCAGGCTTGTATTGTGACATTTTCCAGGGTATACTTTGTTCC
[G,T]
AGTTTCCCTTCCTGCCTTGATGTTGTGATACAGTGTAGGTGACCAGGGAAGCCTATCTGT

FIGURE 3S

AGTTGATGGCAGGTATTACAGTCCCATCACAGGTGGTACAAGATAAAGTAATTTGCTGGG
GCTTAGAGGACTGGTTGAGTACTTCCAGCCTGGGGCATAGGATCCACGCAAGGATTTATA
TAGAAAACATGCCAGGTATGATTAAGGTAGAGGTTGATTTGGAGGACCTTCTTAACCTAA
ATTAATATTTTAATATGTCGGAAGTGTTAGAGACAAGTTTTTGAGCTGGGTTCCTTTTAT

15207 CTTGATGTTGTGATACAGTGTAGGTGACCAGGGAAGCCTATCTGTAGTTGATGGCAGGTA
TTACAGTCCCATCACAGGTGGTACAAGATAAAGTAATTTGCTGGGGCTTAGAGGACTGGT
TGAGTACTTCCAGCCTGGGGCATAGGATCCACGCAAGGATTTATATAGAAAACATGCCAG
GTATGATTAAGGTAGAGGTTGATTTGGAGGACCTTCTTAACCTAAATTAATATTTTAATA
TGTCGGAAGTGTTAGAGACAAGTTTTTGAGCTGGGTTCCTTTTATATTTCTGGTTTGCCC
[C,T]
ACCCTTTTATCTAGTTTGCGCAAGGAACAAAATACATGGAAGTACTTCTACACCTACTGC
ACATATGCATGCACACACCTGGCTCTTCTAGCAAGTCAAGGGCTCAGCAAAAACCCCTAG
TTAGGGGGTGCAAATAGGAACCCCAAACACTTCCATGAGTTTCATGGGTTACTTCCTTTT
ATTTTTTTGAGACAGGGTCTTGCTCTGTTGTCCAGGCTGGAGTGCACTGGCACAATCATG
GCTCACTGCAACCTCCATCTCCTGGGCTCAAGTGATCCTCCCACCTTAGTTTCCTAAGTA

16162 GCAGACTGAAACTGGCTCTCTAAAGGTGAGCTGGAGTAGTCATTTGCAAAATGTGGTCTG
CACACTTTGTGGGCTTCCCAAGACCATTTCAAGAAGTCTATGAGGCTAAAACTCTCTTCA
TAATAATACTAAGATGTTATCTGCTTTTTCACTTGTGGATATTTGCACTTATAATGTAGA
AGCAATGGTGGGTAAAATTACACTGTAGAACGAATCAAGGCAGTGGCACCAAATTATACT
AGTTGTCGTTGTATTTTTCACTGCCACACATGCGCAAAGAAAAAAGCCCTTTGCACTTAA
[G,T]
AATGTCTTTGATGAAACTGTAGGATTACTAATATTTAAAAATTTGAGACCCTTCAGTATA
GGTCTTTAATATTCTGTGTGGCAAAATGGGAAGTATGCATGAAGTACTTCTATGAGTACC
AAAATATGTTACTTGTCTTAAGGCAAAGACCTCGAGTGATTATATGAGTTGTCAACCAAA
CTTGCTGCCTTTTTTTTTTTTTTCATAGAACTAGAAAGAACAACTAACAAACTGTAGGTCA
TTCAGACCCGAGTACTTGTAAGACATTTTCTTGAAAATGAAAGAAATCAGCCCATCACCT

16364 CTGTAGAACGAATCAAGGCAGTGGCACCAAATTATACTAGTTGTCGTTGTATTTTTCACT
GCCACACATGCGCAAAGAAAAAAGCCCTTTGCACTTAAGAATGTCTTTGATGAAACTGTA
GGATTACTAATATTTAAAAATTTGAGACCCTTCAGTATAGGTCTTTAATATTCTGTGTGG
CAAAATGGGAAGTATGCATGAAGTACTTCTATGAGTACCAAAATATGTTACTTGTCTTAA
GGCAAAGACCTCGAGTGATTATATGAGTTGTCAACCAAACTTGCTGCCTTTTTTTTTTT
[T,-]
CATAGAACTAGAAAGAACAACTAACAAACTGTAGGTCATTCAGACCCGAGTACTTGTAAG
ACATTTTCTTGAAAATGAAAGAAATCAGCCCATCACCTCAAGGAAAACAATAGATAATAC
ATATCTGTTGCCCAGAATAAAATTCAAGCTTTCAAGCAAAATTAGGAAAAAAAACCAACT
TGTATCCAGTACCATGAGCTTGATAGCCCCTCTACTTGAAGACTTTTCTGATGAGATTAG
TGGTGATATTAACAAATATGACTTTTTGATATTATTAATATACAATGAAGATGTTAACAT

16411 TGTATTTTTCACTGCCACACATGCGCAAAGAAAAAAGCCCTTTGCACTTAAGAATGTCTT
TGATGAAACTGTAGGATTACTAATATTTAAAAATTTGAGACCCTTCAGTATAGGTCTTTA
ATATTCTGTGTGGCAAAATGGGAAGTATGCATGAAGTACTTCTATGAGTACCAAAATATG
TTACTTGTCTTAAGGCAAAGACCTCGAGTGATTATATGAGTTGTCAACCAAACTTGCTGC
CTTTTTTTTTTTTTCATAGAACTAGAAAGAACAACTAACAAACTGTAGGTCATTCAGACC
[T,C]
GAGTACTTGTAAGACATTTTCTTGAAAATGAAAGAAATCAGCCCATCACCTCAAGGAAAA
CAATAGATAATACATATCTGTTGCCCAGAATAAAATTCAAGCTTTCAAGCAAAATTAGGA
AAAAAAACCAACTTGTATCCAGTACCATGAGCTTGATAGCCCCTCTACTTGAAGACTTTT
CTGATGAGATTAGTGGTGATATTAACAAATATGACTTTTTGATATTATTAATATACAATG
AAGATGTTAACATTTGGAAGATCTGTGTAAACTCAACCAAAGTATGATGTTAGGAATTCT

16636 AACCAAACTTGCTGCCTTTTTTTTTTTTTTCATAGAACTAGAAAGAACAACTAACAAACTG
TAGGTCATTCAGACCCGAGTACTTGTAAGACATTTTCTTGAAAATGAAAGAAATCAGCCC
ATCACCTCAAGGAAAACAATAGATAATACATATCTGTTGCCCAGAATAAAATTCAAGCTT
TCAAGCAAAATTAGGAAAAAAAACCAACTTGTATCCAGTACCATGAGCTTGATAGCCCCT
CTACTTGAAGACTTTTCTGATGAGATTAGTGGTGATATTAACAAATATGACTTTTTGATA
[T,C]
TATTAATATACAATGAAGATGTTAACATTTGGAAGATCTGTGTAAACTCAACCAAAGTAT
GATGTTAGGAATTCTGCATGGGTAAAAGATCCATTGAAAGAGCAAGATCACCAATGGATT
TTTTTTTTCTTTTTTTTTTGAGACAGTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGC
ACAATCTTGGCTCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTTCTGCCTCAGCC
TCCCGAGTAGCTGGGATTACAGGTGCCTGCCACCACTCCCAGCTAATTTTTATATTTTTA

16802 TAAAATTCAAGCTTTCAAGCAAAATTAGGAAAAAAAACCAACTTGTATCCAGTACCATGA
GCTTGATAGCCCCTCTACTTGAAGACTTTTCTGATGAGATTAGTGGTGATATTAACAAAT
ATGACTTTTTGATATTATTAATATACAATGAAGATGTTAACATTTGGAAGATCTGTGTAA

FIGURE 3T

ACTCAACCAAAGTATGATGTTAGGAATTCTGCATGGGTAAAAGATCCATTGAAAGAGCAA
GATCACCAATGGATTTTTTTTTTCTTTTTTTTTTGAGACAGTCTTGCTCTGTCACCCAGG
[C,T]
TGGAGTGCAGTGGCACAATCTTGGCTCACTGCAACCTCTGCCTCCCGGATTCAAGCGATT
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTGCCTGCCACCACTCCCAGCTAA
TTTTTATATTTTTAGTAGAGACGGGGTTTCGCCATGTTGGCCAGGATAGTCTCAATCTCT
TGACCTCATGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCAC
TGCACCTGGCCTGACTTTTTTTTTTTTTTTAAATACTAAATGTATCAGGGACTTCTGGCCT

17111 AGTGGCACAATCTTGGCTCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTTCTGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGCCTGCCACCACTCCCAGCTAATTTTTATA
TTTTTAGTAGAGACGGGGTTTCGCCATGTTGGCCAGGATAGTCTCAATCTCTTGACCTCA
TGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCACCTG
GCCTGACTTTTTTTTTTTTTTTAAATACTAAATGTATCAGGGACTTCTGGCCTCTTATGGT
[A,G]
TGGTGTGACTTTTATGCTGTTCACTTTGTATCTTTCTGTTACAGGGTTTGGGGCTTCTGT
TATTATTATTATTATTTTTTAATTTCCTCTGTTCTCTTACCAGTGTTTGTCCGTCATTGT
TTGGTTTGTCATCCTCTGTTGCAGTTTTGGGATCTGAGTCTTTTTTTTTTTTTTGAGATGG
AGTCTCCCTCTATTGCCTAGGCTGGAGTACAGTGGCACGATCTTAACTCACTGCAACCTC
TGCCTCCCGGGTTCAAGCAATTCTCCTACCTTAACCTCCTGAGAAGCTGGGATTACAGGC

17276 AATCTCTTGACCTCATGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCAT
GAGCCACTGCACCTGGCCTGACTTTTTTTTTTTTTTAAATACTAAATGTATCAGGGACTT
CTGGCCTCTTATGGTGTGGTGTGACTTTTATGCTGTTCACTTTGTATCTTTCTGTTACAG
GGTTTGGGGCTTCTGTTATTATTATTATTATTTTTTAATTTCCTCTGTTCTCTTACCAGT
GTTTGTCCGTCATTGTTTGGTTTGTCATCCTCTGTTGCAGTTTTGGGATCTGAGTCTTTT
[T,G]
TTTTTTTTGAGATGGAGTCTCCCTCTATTGCCTAGGCTGGAGTACAGTGGCACGATCTTA
ACTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAATTCTCCTACCTTAACCTCCTGAGAA
GCTGGGATTACAGGCACATGCCGCTATGCCTGGCTAATTTCTGTATTTTTAGTAGAGACG
GGGTTTCGCCTTGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCGCTT
CGGCCTCCCAAAGTAGTGGGATTATAGGCATGAGCCACTGTGCCTGGCCAGGTCTGAGCC

17372 AAATACTAAATGTATCAGGGACTTCTGGCCTCTTATGGTGTGGTGTGACTTTTATGCTGT
TCACTTTGTATCTTTCTGTTACAGGGTTTGGGGCTTCTGTTATTATTATTATTATTTTTT
AATTTCCTCTGTTCTCTTACCAGTGTTTGTCCGTCATTGTTTGGTTTGTCATCCTCTGTT
GCAGTTTTGGGATCTGAGTCTTTTTTTTTTTTTGAGATGGAGTCTCCCTCTATTGCCTAG
GCTGGAGTACAGTGGCACGATCTTAACTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAA
[C,T]
TCTCCTACCTTAACCTCCTGAGAAGCTGGGATTACAGGCACATGCCGCTATGCCTGGCTA
ATTTCTGTATTTTTAGTAGAGACGGGGTTTCGCCTTGTTGGCCAGGCTGGTCTCGAACTC
CTGACCTCAGGTGATCCACCGCTTCGGCCTCCCAAAGTAGTGGGATTATAGGCATGAGCC
ACTGTGCCTGGCCAGGTCTGAGCCTTTACAGTGGTCAGTTCAGTGGTTAGAACCAGACCC
AAATACACTTGGAAAGGATAGAGTGTCTGAAGAGAGTTGGAGCACCCCTCTGGTCTAATC

18317 AAAAATGCAGGCATCGTGGTAAGGGTCTGCATCAGTGGAGAGGAGTGGTGACAAATTTTA
GGAGGTAGCTTTTTGTTGTTGTTAAAATGTACTTGCTTTAAAACATTTTAAATAGAGAAG
CATTTTAAAAAAATCAGTTGACAAAAAGCGGAATTCAGACATTCATTCACTTAAAGATAT
TTATTGAGAGTGTTCTGTGCGTTAGGCACTGTTCTAAGCTCTTAGAATACATCAGTGAAT
TAAATATTCCTGCCCTCATGGAGCTTACTTCATGGTGGAGAGGATGTACTGAGATGGCTC
[T,G]
AGCAGTTTCTGTCAATAATATGAACTAATGAGTTAGTTACAGATGTCTGCCCATTTTCTA
CAGTCTCCCATGCCCTGTTCCTAAATGGCCAACTGCAAGAATCTTATGTCTTCTTTTTGT
GATTTACCTCCAGTTGACTGCCTGCCCAAAGCCATTCTGGTTTCTTTCGGAGTTGAAGAG
AGACTCAGAGATGTGGGTTGCCCTTAGCTAAGTGCAGTCTTTCTTGATCTGGCATTGCTG
TAAAGATAACTTACCCGTCTCACCTCACATCCCTTAGCCCAGCTCTTCCCACAGTCACAG

18342 TCTGCATCAGTGGAGAGGAGTGGTGACAAATTTTAGGAGGTAGCTTTTTGTTGTTGTTAA
AATGTACTTGCTTTAAAACATTTTAAATAGAGAAGCATTTTAAAAAAATCAGTTGACAAA
AAGCGGAATTCAGACATTCATTCACTTAAAGATATTTATTGAGAGTGTTCTGTGCGTTAG
GCACTGTTCTAAGCTCTTAGAATACATCAGTGAATTAAATATTCCTGCCCTCATGGAGCT
TACTTCATGGTGGAGAGGATGTACTGAGATGGCTCGAGCAGTTTCTGTCAATAATATGAA
[C,T]
TAATGAGTTAGTTACAGATGTCTGCCCATTTTCTACAGTCTCCCATGCCCTGTTCCTAAA
TGGCCAACTGCAAGAATCTTATGTCTTCTTTTTGTGATTTACCTCCAGTTGACTGCCTGC
CCAAAGCCATTCTGGTTTCTTTCGGAGTTGAAGAGAGACTCAGAGATGTGGGTTGCCCTT
AGCTAAGTGCAGTCTTTCTTGATCTGGCATTGCTGTAAAGATAACTTACCCGTCTCACCT
CACATCCCTTAGCCCAGCTCTTCCCACAGTCACAGGAGCCTTCTATTCTGCTGATGTGCA

FIGURE 3U

21828 AACAGTTCTAGGTAAGAAGGGAAGATTTCCTGAAGTAAATTATGTGGTTCCTACCCTCCA
GAGGCTTGTAGTCTGTGTAAGGAAAAAGAAATGTGGGAAGAGAAGCCGGGGAACAAGATA
AGAGACCAGTAGTGGGAGACACCCATAAGAAGAAAGTGTCATGAGCTAGGAGTACACCCT
CAGTGCTCAGAGAGAGAGGAACTTTAAAGATTCTCTTGTCGGCTGTGCCAGATGAGAAAC
GCACATGAGAGATAGGAGCAAAGAAGGCTTCAGGAGAAGGTGAGATAAACTAGAGCAGGG
[C,T]
GTGGAGATGAGTTTGGAGGTGGGAAGTATTTGCAAATTTCTCGTTATGGTAACTCTTCAG
TGTTTGGAGGGAAATATTATGTTTGTTTTCTACATTTAAATGTAGGAAATTGATACTATC
AAGGGCTAAAAATTCTTAAAAAAAAAAAAAAAGAACCACATTAAAACTATGTTCTCTAGA
AAAGTTCCTTTTTGTTGTCATAGAGGAAACTTACTTTCATTCATAGTCACCTTTATCCTG
TGATGCAGATTATATAGTTCTTTTGGCCAAATTATTTTCTGTAACTGGGAGAAGCTAGAT

22674 CATACACATGTTCTTAAAGTTTTTTGCACAGATTGACCATACTATGTATACTGTTTGGAA
ACTTGCTTTTTCCCCTTAAACGTCTGAGACGTTTTTCTCTATCAGCACATAGAGATTTAA
CACATTCTTTTTAACTGCTGTGTAATGTTCCATTTAAGAACGGTCTATAATTTAATCACT
CTGCTTTTGATGATCCTTTAGGTTGTTACCAGCTGCTATTGTTCAACCAGCAGTCTGTTT
TTGGTACATCAGTTTCTGTGTCCTTAATGTGGGACTTGGTTGGTTCTTATATCCAAGTTA
[T,C]
AGAGACAGTGAAGGGGACTATTTCCTTGTGTTTTATGTCAAGGGCTCCCTGTAACTAACA
AAAAAGTGTGAGATGGGATAGGTGGGCAGATGTGTAGAGAGGATGCTAAGGGGCTGGGCA
GTGGTCATGGTGTCTGTGCATGTGTCTCACCTCATGCAGCATTCCAGACGAGAAGCCAGG
AAGGGGACGTCGGAAACCACACAGATAGCACCTCCCTCACCTTCTTCCCAATGCCCCAGA
CCAGTGGCACCTAGCATGGTTTCTTCTCCTGCCAGGGCATCTCGTCCTTGTCACTGCCAG

22683 GTTCTTAAAGTTTTTTGCACAGATTGACCATACTATGTATACTGTTTGGAAACTTGCTTT
TTCCCCTTAAACGTCTGAGACGTTTTTCTCTATCAGCACATAGAGATTTAACACATTCTT
TTTAACTGCTGTGTAATGTTCCATTTAAGAACGGTCTATAATTTAATCACTCTGCTTTTG
ATGATCCTTTAGGTTGTTACCAGCTGCTATTGTTCAACCAGCAGTCTGTTTTTGGTACAT
CAGTTTCTGTGTCCTTAATGTGGGACTTGGTTGGTTCTTATATCCAAGTTATAGAGACAG
[T,C]
GAAGGGGACTATTTCCTTGTGTTTTATGTCAAGGGCTCCCTGTAACTAACAAAAAAGTGT
GAGATGGGATAGGTGGGCAGATGTGTAGAGAGGATGCTAAGGGGCTGGGCAGTGGTCATG
GTGTCTGTGCATGTGTCTCACCTCATGCAGCATTCCAGACGAGAAGCCAGGAAGGGGACG
TCGGAAACCACACAGATAGCACCTCCCTCACCTTCTTCCCAATGCCCCAGACCAGTGGCA
CCTAGCATGGTTTCTTCTCCTGCCAGGGCATCTCGTCCTTGTCACTGCCAGGAAGGGTCT

22822 TCCATTTAAGAACGGTCTATAATTTAATCACTCTGCTTTTGATGATCCTTTAGGTTGTTA
CCAGCTGCTATTGTTCAACCAGCAGTCTGTTTTTGGTACATCAGTTTCTGTGTCCTTAAT
GTGGGACTTGGTTGGTTCTTATATCCAAGTTATAGAGACAGTGAAGGGGACTATTTCCTT
GTGTTTTATGTCAAGGGCTCCCTGTAACTAACAAAAAAGTGTGAGATGGGATAGGTGGGC
AGATGTGTAGAGAGGATGCTAAGGGGCTGGGCAGTGGTCATGGTGTCTGTGCATGTGTCT
[C,G]
ACCTCATGCAGCATTCCAGACGAGAAGCCAGGAAGGGGACGTCGGAAACCACACAGATAG
CACCTCCCTCACCTTCTTCCCAATGCCCCAGACCAGTGGCACCTAGCATGGTTTCTTCTC
CTGCCAGGGCATCTCGTCCTTGTCACTGCCAGGAAGGGTCTGTGATGGCTTGGGGAAAAG
CACTGTTAAAAAAACACTTAATGGGCACAATGTACACTGTTTGGGTGATGGGTACACTAA
ACGCCCAGGCACTACCACTATGCAGTATATCCATTTAACAAAACAGCACTTGTACTCCCT

23073 GAGGATGCTAAGGGGCTGGGCAGTGGTCATGGTGTCTGTGCATGTGTCTCACCTCATGCA
GCATTCCAGACGAGAAGCCAGGAAGGGGACGTCGGAAACCACACAGATAGCACCTCCCTC
ACCTTCTTCCCAATGCCCCAGACCAGTGGCACCTAGCATGGTTTCTTCTCCTGCCAGGGC
ATCTCGTCCTTGTCACTGCCAGGAAGGGTCTGTGATGGCTTGGGGAAAAGCACTGTTAAA
AAAACACTTAATGGGCACAATGTACACTGTTTGGGTGATGGGTACACTAAACGCCCAGGC
[G,A]
CTACCACTATGCAGTATATCCATTTAACAAAACAGCACTTGTACTCCCTAAATCTATTAA
AAAACAAAAACAAAAAACACCTCCCCTTCTGGGAGCATTGCATTTGTATTGTAACAGTCT
TTGTATTCCTTCCTTCCCCACCTCCAGACGTGTTGTGGACTTCTTCCTGATTGTCACCCA
GCTGGGATTCTGCTGTGTCTATTTTGTGTTTCTGGCTGACAACTTTAAACAGGTAGGCAC
CTGGTTAAAAAAGAAAAAAAAAAAAAAAAACCAGAGCGAGAATGGCAAAAGATGATTGAAG

23343 TTGGGTGATGGGTACACTAAACGCCCAGGCACTACCACTATGCAGTATATCCATTTAACA
AAACAGCACTTGTACTCCCTAAATCTATTAAAAAACAAAAACAAAAAACACCTCCCCTTC
TGGGAGCATTGCATTTGTATTGTAACAGTCTTTGTATTCCTTCCTTCCCCACCTCCAGAC
GTGTTGTGGACTTCTTCCTGATTGTCACCCAGCTGGGATTCTGCTGTGTCTATTTTGTGT
TTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAAGAAAAAAAAAAAAAAAA
[C,-]
CAGAGCGAGAATGGCAAAAGATGATTGAAGTTTTTGTTTAGGATTTTTTCCAAATCAGCT

FIGURE 3V

TTTGTCAACAAAAGAGTTAAAGTTTTCATATTTTACATAGATCTACGTCTTCTATTTGAT
TCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTCTTGTCGACCCTCCTGTCC
TAGGTACATATGATCAAACCTAGCTCAGACAATTGGGTTGCTGATGATAGTCGTGAAGTT
CTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGGCCTTGTTCACACACCTGAGCCTT

23396 TTTAACAAAACAGCACTTGTACTCCCTAAATCTATTAAAAAACAAAAACAAAAAACACCT
CCCCTTCTGGGAGCATTGCATTTGTATTGTAACAGTCTTTGTATTCCTTCCTTCCCCACC
TCCAGACGTGTTGTGGACTTCTTCCTGATTGTCACCCAGCTGGGATTCTGCTGTGTCTAT
TTTGTGTTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAAGAAAAAAAAA
AAAAAAACCAGAGCGAGAATGGCAAAAGATGATTGAAGTTTTTGTTTAGGATTTTTTCCA
[T,A]
ATCAGCTTTTGTCAACAAAAGAGTTAAAGTTTTCATATTTTACATAGATCTACGTCTTCT
ATTTGATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTCTTGTCGACCCT
CCTGTCCTAGGTACATATGATCAAACCTAGCTCAGACAATTGGGTTGCTGATGATAGTCG
TGAAGTTCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGGCCTTGTTCACACACCT
GAGCCTTTCCTCAGGAACCTCTTCCAGCAGAGGATCCACCGGCCTCTGTTGTTTGAGAGG

23511 CCACCTCCAGACGTGTTGTGGACTTCTTCCTGATTGTCACCCAGCTGGGATTCTGCTGTG
TCTATTTTGTGTTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAAGAAAA
AAAAAAAAAAAACCAGAGCGAGAATGGCAAAAGATGATTGAAGTTTTTGTTTAGGATTTT
TTCCAAATCAGCTTTTGTCAACAAAAGAGTTAAAGTTTTCATATTTTACATAGATCTACG
TCTTCTATTTGATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTCTTGTC
[G,C]
ACCCTCCTGTCCTAGGTACATATGATCAAACCTAGCTCAGACAATTGGGTTGCTGATGAT
AGTCGTGAAGTTCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGGCCTTGTTCACA
CACCTGAGCCTTTCCTCAGGAACCTCTTCCAGCAGAGGATCCACCGGCCTCTGTTGTTTG
AGAGGTGTTTCCGTTTTCTTCCTTCCCCTCATTCTAGGTGATAGAAGCGGCCAATGGGAC
CACCAATAACTGCCACAACAATGAGACGGTGATTCTGACGCCTACCATGGACTCGCGACT

23522 CGTGTTGTGGACTTCTTCCTGATTGTCACCCAGCTGGGATTCTGCTGTGTCTATTTTGTG
TTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAAGAAAAAAAAAAAAAAA
ACCAGAGCGAGAATGGCAAAAGATGATTGAAGTTTTTGTTTAGGATTTTTTCCAAATCAG
CTTTTGTCAACAAAAGAGTTAAAGTTTTCATATTTTACATAGATCTACGTCTTCTATTTG
ATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTCTTGTCGACCCTCCTGT
[C,G]
CTAGGTACATATGATCAAACCTAGCTCAGACAATTGGGTTGCTGATGATAGTCGTGAAGT
TCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGGCCTTGTTCACACACCTGAGCCT
TTCCTCAGGAACCTCTTCCAGCAGAGGATCCACCGGCCTCTGTTGTTTGAGAGGTGTTTC
CGTTTTCTTCCTTCCCCTCATTCTAGGTGATAGAAGCGGCCAATGGGACCACCAATAACT
GCCACAACAATGAGACGGTGATTCTGACGCCTACCATGGACTCGCGACTCTACATGCTCT

23582 TTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAAGAAAAAAAAAAAAAAA
ACCAGAGCGAGAATGGCAAAAGATGATTGAAGTTTTTGTTTAGGATTTTTTCCAAATCAG
CTTTTGTCAACAAAAGAGTTAAAGTTTTCATATTTTACATAGATCTACGTCTTCTATTTG
ATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTCTTGTCGACCCTCCTGT
CCTAGGTACATATGATCAAACCTAGCTCAGACAATTGGGTTGCTGATGATAGTCGTGAAG
[T,A]
TCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGGCCTTGTTCACACACCTGAGCCT
TTCCTCAGGAACCTCTTCCAGCAGAGGATCCACCGGCCTCTGTTGTTTGAGAGGTGTTTC
CGTTTTCTTCCTTCCCCTCATTCTAGGTGATAGAAGCGGCCAATGGGACCACCAATAACT
GCCACAACAATGAGACGGTGATTCTGACGCCTACCATGGACTCGCGACTCTACATGCTCT
CCTTCCTGCCCTTCCTGGTGCTGCTGGTTTTCATCAGGAACCTCCGAGCCCTGTCCATCT

24977 TCCATGAGTTATGACAACCAAAAATGTCTCCAGACATTGCCAAACCTTCCTGGGGGGCAA
AATCGCCCCCCCACCCAGGGGGCACTGGTTTAGACTTTTTTCAATTAGATGGTTAATTCA
TGATCATTGTATACAGTTGGAAAATAGAGGAAAATGTTAAGATTAAAATAAAAAATAATT
TTTCTAACCTGTATTTAGATAAGTAATTCCTTATCAACTCCAGTTAATTTTTATTTGTCA
AAATTATAAATTCACTTGTTCCTTGCCCTCACTTAACCCATGCAGGCAAGTCTGTGGGGT
[G,A]
GCATGAGAGAGAACATCTGTATACAGATGGGTAGAAAATCAGGCTGAGAAAAATGTGCCC
TTAAACACTATGGCTGTTTGTGAAAATGAGAATGCATTTTCTAAGGCTTGAGAAAAGGAA
AAAAGTAAAAGCGGGTAAATAAAAGCATAACTTAAAAAAAAAAATACTTAAATTCAGTTC
CCCAAATAATTCATCAGTACATATTCATTAAAATGCAGACAACACAAATACCTCTTGAAT
ACCATGTCCCCACCCCGAGTCTCCTCTCAGGGACCCGCTGTATGTGATTGGTCTGTCTCA

25131 TGTTAAGATTAAAATAAAAAATAATTTTTCTAACCTGTATTTAGATAAGTAATTCCTTAT
CAACTCCAGTTAATTTTTATTTGTCAAAATTATAAATTCACTTGTTCCTTGCCCTCACTT
AACCCATGCAGGCAAGTCTGTGGGGTGGCATGAGAGAGAACATCTGTATACAGATGGGTA

FIGURE 3W

GAAAATCAGGCTGAGAAAAATGTGCCCTTAAACACTATGGCTGTTTGTGAAAATGAGAAT
GCATTTTCTAAGGCTTGAGAAAAGGAAAAAAGTAAAAGCGGGTAAATAAAAGCATAACTT
[A,-]
AAAAAAAAAATACTTAAATTCAGTTCCCCAAATAATTCATCAGTACATATTCATTAAAAT
GCAGACAACACAAATACCTCTTGAATACCATGTCCCCACCCCGAGTCTCCTCTCAGGGAC
CCGCTGTATGTGATTGGTCTGTCTCATTCTAGATCCTGTGAATGGATTTACAGCCCATGT
AAGTATATTGAGAAATACATTGAAATATATTTTGTTTTCATTTTTGAAACATAATTTTTT
AAAGTTACATGTTCATCTACCTTGCTTTTTTCCACCTTAAAAATGCCTTAGTGAGCCTTC

25178 AGTAATTCCTTATCAACTCCAGTTAATTTTTATTTGTCAAAATTATAAATTCACTTGTTC
CTTGCCCTCACTTAACCCATGCAGGCAAGTCTGTGGGGTGGCATGAGAGAGAACATCTGT
ATACAGATGGGTAGAAAATCAGGCTGAGAAAAATGTGCCCTTAAACACTATGGCTGTTTG
TGAAAATGAGAATGCATTTTCTAAGGCTTGAGAAAAGGAAAAAAGTAAAAGCGGGTAAAT
AAAAGCATAACTTAAAAAAAAAAATACTTAAATTCAGTTCCCCAAATAATTCATCAGTAC
[G,A]
TATTCATTAAAATGCAGACAACACAAATACCTCTTGAATACCATGTCCCCACCCCGAGTC
TCCTCTCAGGGACCCGCTGTATGTGATTGGTCTGTCTCATTCTAGATCCTGTGAATGGAT
TTACAGCCCATGTAAGTATATTGAGAAATACATTGAAATATATTTTGTTTTCATTTTTGA
AACATAATTTTTTTAAAGTTACATGTTCATCTACCTTGCTTTTTTCCACCTTAAAAATGCC
TTAGTGAGCCTTCCAGGTTAGTATTCCTGGCTCTACCTTGTTGTTGTTAGTTGTCACATT

25351 CTGTTTGTGAAAATGAGAATGCATTTTCTAAGGCTTGAGAAAAGGAAAAAAGTAAAAGCG
GGTAAATAAAAGCATAACTTAAAAAAAAAAATACTTAAATTCAGTTCCCCAAATAATTCA
TCAGTACATATTCATTAAAATGCAGACAACACAAATACCTCTTGAATACCATGTCCCCAC
CCCGAGTCTCCTCTCAGGGACCCGCTGTATGTGATTGGTCTGTCTCATTCTAGATCCTGT
GAATGGATTTACAGCCCATGTAAGTATATTGAGAAATACATTGAAATATATTTTGTTTTC
[G,A]
TTTTTGAAACATAATTTTTTTAAAGTTACATGTTCATCTACCTTGCTTTTTTCCACCTTAA
AAATGCCTTAGTGAGCCTTCCAGGTTAGTATTCCTGGCTCTACCTTGTTGTTGTTAGTTG
TCACATTGTATCACAGCAAGGAGATTTGCTGCCATTTATTTAACAAGTCCTCACTCAGTG
GCTATCAGGCCATGGATAATTTTTAGTATTATTTCAGTATTAAGACAGTGAGATGTTCTT
ATACATTCCTTTTTGTGGACTTGTATAAATACTTAAGATATTTGTCTAGATGTGTAATTG

25380 AAGGCTTGAGAAAAGGAAAAAAGTAAAAGCGGGTAAATAAAAGCATAACTTAAAAAAAAA
AATACTTAAATTCAGTTCCCCAAATAATTCATCAGTACATATTCATTAAAATGCAGACAA
CACAAATACCTCTTGAATACCATGTCCCCACCCCGAGTCTCCTCTCAGGGACCCGCTGTA
TGTGATTGGTCTGTCTCATTCTAGATCCTGTGAATGGATTTACAGCCCATGTAAGTATAT
TGAGAAATACATTGAAATATATTTTGTTTTCATTTTTGAAACATAATTTTTTTAAAGTTAC
[A,G]
TGTTCATCTACCTTGCTTTTTTTCCACCTTAAAAATGCCTTAGTGAGCCTTCCAGGTTAGT
ATTCCTGGCTCTACCTTGTTGTTGTTAGTTGTCACATTGTATCACAGCAAGGAGATTTGC
TGCCATTTATTTAACAAGTCCTCACTCAGTGGCTATCAGGCCATGGATAATTTTTAGTAT
TATTTCAGTATTAAGACAGTGAGATGTTCTTATACATTCCTTTTTGTGGACTTGTATAAA
TACTTAAGATATTTGTCTAGATGTGTAATTGCTGAAGAGTGTGCACTTTTGAATTTTTGT

26737 GTATCCATCCAGAGTGATGTCTATGCATAGTACAACCAGGACACAGAGCAATGTCTGCAT
AAGGGCAGCCCTGCTGATTTCTTGAGAGCAATTCTGAGTCTTCCTCTGGGCTTAGCCAGA
AGTTGTGCTGTGATCAAATAGTGCCGTCTGCCTGGAGTACAGCATGGGGGAAGAGGTTTG
GCTGTGTTTTGATGTAGTCACTGCCCATAGTGTTGTAGTTGCTTCATTTTGATGTGTCAT
ACAGCTAAAGATGCTCCCTTTAGGTCATTTTTGTTGCCGCTGCCTCTGCGGCTTGTTACT
[G,A]
CTGTTCTGTTTTGGCATTGTGCCCCACTTACCATGAGGATTCCCCTACTGTTCAATGTTT
CTGAATTTTTTCCCTAATCCTAAGCATGTACATGACTGTTCCTCTTGCCCCTCATGCACG
TGCCATTGTAGGTAGCAGACCAAGGTCTTCCACAGAGAGCAGGTTCCTCTCTGTCTTCAG
CATGTGGAGTCTCAAATGGAACAGTTCTGGGCAGAGTGCTTTGCACAGAGGGTGCTCCCA
ATAAATGTTTTATCACTGCATATCGTTGCTTCTGAGATGTATTTTTTCATAGTTATAACA

26829 TCTGAGTCTTCCTCTGGGCTTAGCCAGAAGTTGTGCTGTGATCAAATAGTGCCGTCTGCC
TGGAGTACAGCATGGGGGAAGAGGTTTGGCTGTGTTTTGATGTAGTCACTGCCCATAGTG
TTGTAGTTGCTTCATTTTGATGTGTCATACAGCTAAAGATGCTCCCTTTAGGTCATTTTT
GTTGCCGCTGCCTCTGCGGCTTGTTACTACTGTTCTGTTTTGGCATTGTGCCCCACTTAC
CATGAGGATTCCCCTACTGTTCAATGTTTCTGAATTTTTTCCCTAATCCTAAGCATGTAC
[G,A]
TGACTGTTCCTCTTGCCCCTCATGCACGTGCCATTGTAGGTAGCAGACCAAGGTCTTCCA
CAGAGAGCAGGTTCCTCTCTGTCTTCAGCATGTGGAGTCTCAAATGGAACAGTTCTGGGC
AGAGTGCTTTGCACAGAGGGTGCTCCCAATAAATGTTTTATCACTGCATATCGTTGCTTC
TGAGATGTATTTTTTCATAGTTATAACAGTTTCAGGATTGCAAGAGTACATCTCACAATC
CATGTGTACCTTTAACAGCATTTTCTCAAAATACTGTTATTATAATTGATAATATGGTAA

FIGURE 3X

27423 ATGGTAAGACCTCACTTAATATCATTGATACATTCTTAGAAACTGCAATACATTAAATGT
ATGTATAGCGAAATCAGTTTTTTTCTCATCAATGTTATAACAAAACAGCGTTGAAGGAAG
TGACTGTACGTCATTTCACTTAAAGTCTCAGTTTCCAAGAACTTATTGACGACAAGGGAG
GACTTACTGTGTTGTAGAATTGAGGAGATATGTTAATAACGAGCTGATTTTAACATGTAT
GTTTCTTTATAAATTAAACTTTTCTCATTTAGTTGGTTGGGTCAGTAGCAATCAGTAAGT
[G,A]
TGTAGAATAATACACTTCTTCTGCTGGCCTCATTCCCACAATATCCCCACATATGGATTG
TGAAATTCCCAGTCTGATACTTGAATCTGATCTGATGTATGAATAAGAGCAGGAGTCATT
CACTAACCAACAGATAGCACCTGTTTCCAATAACTTAGGTTACATTTGTGACTCAGGAAT
AATTACAGGCCACTCTTGCTCTCAAGTCCCATTGTAAAGGAAAAATACCTATTACCCTGT
CTTCATTCCAGGTATTGAAATGCTTCTTACAAAGGGATCTAACAGATTTCTTAGCAGGGG

27735 ACACTTCTTCTGCTGGCCTCATTCCCACAATATCCCCACATATGGATTGTGAAATTCCCA
GTCTGATACTTGAATCTGATCTGATGTATGAATAAGAGCAGGAGTCATTCACTAACCAAC
AGATAGCACCTGTTTCCAATAACTTAGGTTACATTTGTGACTCAGGAATAATTACAGGCC
ACTCTTGCTCTCAAGTCCCATTGTAAAGGAAAAATACCTATTACCCTGTCTTCATTCCAG
GTATTGAAATGCTTCTTACAAAGGGATCTAACAGATTTCTTAGCAGGGGCCCAGGGAAAC
[G,A]
CATTTATTTAATTTTTTTATTTTTTCAAAAGCAATATTACTGCTTTGAAATCTTTCAAAG
TGAAGGCTGTTATAGAGCTTAATAATGGATCTCCTTTTACTTGCCTGAAATTATTCTGAA
GCCTGTTAAGAGCATGCCCCGTATTATCCAAATAGCCATACAGTTAAATCAATTTTAAAA
CATTGTAAAAGGCTGTTTTAACATCAATTTTTATTTTAATTGAAGCAACATACACATGTG
GTTTAGAAAACCAAATTGTAAAAAGACAGCAGCTTTGAATCCCTCCTCCCCACCCTGCCC

29875 GTATTAGACTTTCTTCCTTTAGAGAATCTTGGATTTGTTAAAAGGTATGACCTCTCCGAT
TCAGAGTTCAAATCTTGAATTTCTGTATAGCCTTTTGCTTTGTTTTGCTTTCTGTCTTTC
AGAGGATCCCAGACCCCAGCCACCTCCCCTTGGTGGCCCCTTGGAAGACCTACCCTCTCT
TCTTTGGCACAGCGATTTTTTCATTTGAAGGCATTGGAATGGTAAGAGCTGCACTGTGAT
TTGGGCTAGTGTTCTCTGGTGCCCTTGGTGTTCTCCAGGTCTGTTTCAAGGAATGCTGAG
[C,G]
AAACATTGTTAGAAAGTATCTTCTGAGGCCAGGCATGGTGGCTCACGCCTGTAATCTCAG
CACTTTGGGAGGCCTAGACTGGTGGATCACTTGAGGTCAGGAGTTCGAAACCAGCCTGGC
CAACATGGTGAAACCCCATCTCTACTAAATATACAAAAATCAGCTAGGCATGGTGGCACA
CGCCTATAATCCCAGCCACTCGAGAGGCTGAGGCAGGAGAATTGCTTGAACTGGGGAGAC
GGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGAC

30356 CGCCTATAATCCCAGCCACTCGAGAGGCTGAGGCAGGAGAATTGCTTGAACTGGGGAGAC
GGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGAC
TCTGTCTCAAAAAAAAAAAAAAAAAAAAGAAATTATCTTCTGTAACTCACTGGTCAGTTA
GTGAATAGTGTTTCGGGGATTCCATTGAGATTTCCCAGCTTCAACTTTTCAAGACAAATT
ATATGTAATTTTAAAATGTTTACATTCAAGGCCCCTTCACTGCACACTCATCTCCTATGT
[G,T]
TGCAGTAAGGAATAGCATATGGCAATCAGGAAGGCAGGGTCTAGAGTCAGACTGACATGG
GGGTAAGTCCTGGCTCTGCCATAGAGTAGCTGTGTGACCTTGAGCAAGGGCTTCATCTCT
TTGAGCCTTCATTATTTGTTCCTGAAAAGTGAGCTTAATGATTCGTAGTTATTAGGATTA
AATGAGATATGTGCAAAATGCTTTGCACAGACCCTGACACATGGTAAATGTTTAATAGAT
TTTTATTTTATTAATAATGTTATTTTATTATTGAATCAATAAATGCATGAATAATTTCTC

31344 GTGATCACATGTAAATCACTGGTCTCCCTGAGGTTTTATACCTTGCCCTGTGCTCTATCT
TAGGGCTTTTCATTGCCCATGAAGAGTGTCTCACTGTAATTCAGAAACACAAATGGTTTC
CTGCTGCTGGGGCAGAGGCCTGAGTGGGCCCATACTTCAGCAGTGAGAAAGAGATCCCAA
GAACTCAGGACTGGAAAGAAGAGGCTGAGAAAGTGTGGAAAGATGCCCAGAGACCTTAGG
TTCTTGGGCATCCTAAGGGACCTTGTGCTAAATTTTTAGTAGCTTTCCCTAACAGCACAG
[C,T]
GCAGAAATTGTTTGCTTGGTTTTATTACCCAAGACTTGTACACAAAGTTATTCTGCAAAC
ATCATTTGTTTTCAAGATTTCTTTGTATTTCTATTTTTTTACAATAGAGAGAGAACACTG
CTAGATTGACTCTTAGTTTTGGATCTAGGGCTTGTTCATTGCATCGGGGTAAAGTGCCAG
GCTGCACACTGTATTCACCGTGTGCTCTGTGTTCATGCAGCTGTCACAGGCCAGATATGG
GCTCCCTGCCCTCTGGCTCTTTGTATTCTTGGTATGATGGAAACTGACAGATACATTTAG

32570 GGGCCTTCTGGTTGCCATGTGGCCTCACTCACAGGGCAAGGGACATTCCCAGCCTGCAGG
GATCCTGCAGCAGGAGGACAGAGCACTGGCCTGAGCCAGGAGTCCTGGGCTCGTGTCCTG
TCCACCCTTACTTTTAGGACTTCCTAGCTAGGCAGTGGGCTGCAGAGTCCCTTCTAGTCC
CAAGAGCATAACGTCTGATGAAATAACTTTATTTAAAGAGCAGATGTGCTTCTGGAGAAT
TCTGGGGATAAAAGAGTTACTTTTTTTCTGAGGTTTTTTTTTTTCTTGGCCATTAACTTT
[T,G]
CTTTTTTCTGCACTTTCTCTCCTCTCTCACTACTCTCTCATAGGTTCTGCCCCTGGAAAA

FIGURE 3Y

CAAAATGAAGGATCCTCGGAAGTTCCCACTCATCCTGTACCTGGGCATGGTCATCGTCAC
CATCCTCTACATCAGCCTGGGGTGTCTGGGGTACCTGCAATTTGGAGCTAATATCCAAGG
CAGCATAACCCTCAACCTGCCCAACTGCTGGTACGTGGAGGGAGGATGGAAACCTAGGAG
CACTGGATATTTTTAAAAACTAATGGGTCACAGTGTGGATTCTCCCTCTTACTTATCTCT

33220 CTCACTGGCTGCCCTGGACTGCATTCTGTTTGGGGAATTCATGTAGAGCCTTCTGCTGAA
GCCATTGGTGCTGATCAGCCGATGGGTAAGCCATTTCTCCTTGGAATTCCTAAGCTCAGA
AGGACCGAGTATCTAGTCCATTCATGGTAAACCATTCCAAATAGACAGGGAGATGGGAGG
GCAAACCTGCATTTGATTCCCAGCATCGGTTGTGCCTCTCCCTTGGTAGTAACAGGCTTG
ATATGCAGATGGGAGCATCTCACTGTGAGCCGGGGATTGTTGGGAGTCCTTTTGTACCTC
[T,C]
CTTGCATTGGTGAATGTATTATAGGGAAATAGTGAGCCATTTTGAAATGCTTCCTGAAAG
GGTGAATGTCCCAGGGCATGTGCAGAGCAACCATCCTGTTTTGAAGATGAATCATCTCAT
GGTGGAGAGCAGCTGTTAGCAGACACTGAGAAGCTTGTTGAGTGCTCTGCGGATCAGAAT
CAGCTTTCAGTCTAGGCTGGCTGATCTGCCTGGGTGTGCTTTTTATTTTGTTTTGTATTG
TTTTATTTTATTGTATTTTTTAAGACAACAGCACTCAGTATTTCCAGGGGCTTTCCCGTT

33525 CATTGGTGAATGTATTATAGGGAAATAGTGAGCCATTTTGAAATGCTTCCTGAAAGGGTG
AATGTCCCAGGGCATGTGCAGAGCAACCATCCTGTTTTGAAGATGAATCATCTCATGGTG
GAGAGCAGCTGTTAGCAGACACTGAGAAGCTTGTTGAGTGCTCTGCGGATCAGAATCAGC
TTTCAGTCTAGGCTGGCTGATCTGCCTGGGTGTGCTTTTTATTTTGTTTTGTATTGTTTT
ATTTTATTGTATTTTTTAAGACAACAGCACTCAGTATTTCCAGGGGCTTTCCCGTTCAAG
[T,G]
ACGAACCAGGCTTGACCCTGCTTAGCTTCCAAGATCAGGTGAAATTGAGCACATTCAGAA
TGGTATGGCTATAGACCTGGATTCGCTTTTTATTTTTTTATATTCTTTTTCAGTTGATTT
TAACTCGTGAGGCATACCAATTATATATATGGATGCAGTATGTGTGACATTTGGATACAT
ATGTACAATGTGTAATTATCAAATCAGGGTAATTGGCATATCCATCTTGTGTCTACTTTT
AAATTTCCAAATGTTTCTGCCCTTCCAAGAAGGAAGAGGCAGGTGGTAGCTTGGTGTAAC

34589 CATGAGGGTCTGGTTTAAGGTGGAGCTTTGCTTAGGGACAGAGACCTTTCCTTTTAATGA
CCAGGTCAGATCTGTAAGTTGATCACAGACTGTTTTCCTACTCTGTGCAGTCAAGGCACT
GGAGTAATAAAATAGGGATATCCTGTGGTGAGTTACGTCATTTTTGGAAGCTACACTTGA
AGCAGTAGTAGGAAGAGAGCCATAGTGGTATGGAAAGATGGAATTCTGCTCTGGCCTCTT
GGTCCTGCAGTGTCTTCATCTAATTCTAGGGACACTGACTTGGATGGGACAGATATAAAT
[A,G]
GGCTTGTGACATTTTAATTGCAATTTTGTTTTTATTTTTGAAGGCATGTACACCTGTATG
CCCATGGCAAAGATTGAGATTTTCAAAAGGTATATAGAGAGCATTAAGCTTCCACCCCGC
GCCCTCCACTCTAGTTCCCAATTTTACAATTTCCCATTTCAGAGGCAACCATATTCCCAG
TTTCTTTTTTGTTTGTTTGTTTGTTTTGAGATGTTTAGTGTATGATTGTCATGTGGGGTG
AGTGTGTGTTTTTTCCTCCTCTTTTTTCTTTTTTAAGACAAATTGTAGCACTCTGTAGGT

34832 CCTGCAGTGTCTTCATCTAATTCTAGGGACACTGACTTGGATGGGACAGATATAAATAGG
CTTGTGACATTTTAATTGCAATTTTGTTTTTATTTTTGAAGGCATGTACACCTGTATGCC
CATGGCAAAGATTGAGATTTTCAAAAGGTATATAGAGAGCATTAAGCTTCCACCCCGCGC
CCTCCACTCTAGTTCCCAATTTTACAATTTCCCATTTCAGAGGCAACCATATTCCCAGTT
TCTTTTTTGTTTGTTTGTTTGTTTTGAGATGTTTAGTGTATGATTGTCATGTGGGGTGAG
[G,T]
GTGTGTTTTTTCCTCCTCTTTTTTCTTTTTTAAGACAAATTGTAGCACTCTGTAGGTACT
GTATTGCTTCATGCTTTTTTCACTTAAAAAAAGTGATATAAAACTGTCCCCATGATAGTG
ATATGCTATATCATGTGATAGAGTGATATATCATGGGGATAGTTTCATATCACACCATCA
CACCTAGAGTTCTGCCTCATACTTTGTTAAAAGCTATACGGGGGCACCACGATTTACCTA
TCGAGTTCCCACTGGTTAACATTTAAATTGTTTTCAGTCTTTCCTTCTTAAATAATGCTG

35188 GTACTGTATTGCTTCATGCTTTTTTCACTTAAAAAAAGTGATATAAAACTGTCCCCATGA
TAGTGATATGCTATATCATGTGATAGAGTGATATATCATGGGGATAGTTTCATATCACAC
CATCACACCTAGAGTTCTGCCTCATACTTTGTTAAAAGCTATACGGGGGCACCACGATTT
ACCTATCGAGTTCCCACTGGTTAACATTTAAATTGTTTTCAGTCTTTCCTTCTTAAATAA
TGCTGCAGTGAGATATTTTGAATATAAGCTTTTGTGTATGTGTGTGAGGATATCTGTGAG
[A,G]
TAAATTTCTAGACATGAAATTGCTGGGTCCGAAGGACATGTGGGTTTGTATCCTTGATAA
GTGTCAACAAATCGCAATGGGACCATTTTGCACTCTTGCTGATGATGTATAAGTGTGCTG
AGCAGGCTTGGAATGTCTCCTGTCTGTTTCGGCAGGTTGTACCAGTCAGTTAAGCTGCTG
TACTCCATCGGGATCTTTTTCACCTACGCACTCCAGTTCTACGTCCCGGCTGAGATCATC
ATCCCCTTCTTTGTGTCCCGAGCGCCCGAGCACTGTGAGTTAGTGGTGGACCTGTTTGTG

35614 GCTTGGAATGTCTCCTGTCTGTTTCGGCAGGTTGTACCAGTCAGTTAAGCTGCTGTACTC
CATCGGGATCTTTTTCACCTACGCACTCCAGTTCTACGTCCCGGCTGAGATCATCATCCC
CTTCTTTGTGTCCCGAGCGCCCGAGCACTGTGAGTTAGTGGTGGACCTGTTTGTGCGCAC

FIGURE 3Z

AGTGCTGGTCTGCCTGACATGTGAGTAGAAGATGATAATTGCCTTGCTTGTTTTTCCCTA
AAGGGCACCCAGTCTGCAGGCTTTCATGAGAAAAGACAATGTGTGTTGTAGTGAAGCTGG
[G,C]
TATGTTTGTGACAGAGAACCTGGCCCATGGCCTCACTTTCAGAGTTGAGGCACCTCCAGA
TGGGGAAGTGAATTAATTACATATGTACTGTAAAGAACATGGGAATGAGGACAGTGGTTT
ATGTATAGATAGGGTATGAAATGCTGTGGAGGTGGTTATCATTCAGAGTAAAGACATGCG
ATTACTATCCCATATTAAATAAGGTAAAGGTCTGAAAGCCATTTAACCCATATCTGTAAT
GAGTATAAGTTACTCTGATGAAGGGTACTTATTTGCTTTTTCAAATAGTTGTTTTTCCAC

37852 TTTTAACAGCAGCTTCTTTTATTACTATTTCTGGTGTTTTCCTATCTTTTCCCCAACTTT
TCCTCCTCCTCTTCACCCTCCAAAGGGAACAGGAGGAATTCAGTGTAGTTTCTTTTTTTT
TTTCCCTCTTGGAATTCAACTTTCTCACCACTCTCCCCCATCCTCCAAAGATTACTATGG
CTGATACGGACTTTGTGATGCTTAATTTCAAACAGTTGGAGAAGAGGGGGAGGGAAAACA
AGTATTTCATAGGATAGTGCTCATTTTGTTATGATTTCATATCGGACAGTATCTACTTCC
[C,A]
GCCCATATTTTTGGAAATGCGGACTTAGCAGGTCACCTTATGTCCAGACCTTGTGTGGAA
GAGGCTGGCCCCACCTGTGGAGTCTGGAGTTGTAGGATCAACGGTTTTTTAGATTTCTTT
GGAGCAATAACCCATCCATCCTTCAGTGATTCATACTGATTCTCTGTGTCATTTGCCATG
TGAAACATTTTACTTCAGTTTGCTATGAAAATTTCAGAAACCTATTTCTGAAGATATAAT
TACCTAAAATCGCATCATCCAAGAAGCCTGTTCAGACTGGAATGCAGAGCTGCAAAACAT

38643 GTCCTCAACTGTAAGATAGGAAGGGTGTCTGACCTCTAAGGTTTCTCTCAACTCCAAAAT
TCTGTGATTCTGTATAGGTGCTTTGCGCTTGATTTTAAGTTTCTACACAAATATTACTCT
AAAAAAAGAAAGTCAATGTAAAAACATTTGGGAATAAAAGAAGAAATTCCAGTATTCCAC
CAATTTAACAAAGTAATTTTTTTTTTGCATTGTATCTTCTGTGTCTTAATCCTCATGGGTG
CCTTGTAAAAATAGTTGCAATTGTAGTTTACACATAATTTTGTCTTTCACATTTTATTTA
[G,A]
TTTTATATCACAAATATTCATATCTTTCACTAATATTTTCATGACCTCGTGGTATTCCAC
TGTATTGGTGGATCATATTAACTAAGGTACTCCTTTCATGTTGGACATGGTGGTTGTTTC
CCTTGTTTTCGTATTTTTTAAATTTATACCCCCACTAAGTCAAACTTTGTATACTGTCCA
AGACTACTGTGAATTTTAAAGGCATATTTATAGACATTTAAAAGTAACATGGTGAAACCC
CGTCTCTACTAAAAATACAAAAACAAAATTAGCCTGGTGTGGTGGCAGGTGCCTGTAGTC

39198 ATACAAAAACAAAATTAGCCTGGTGTGGTGGCAGGTGCCTGTAGTCCCAGCTACTTGGGA
GGCTGAGGCAGGAGAATGGCATGAACCCAGGAGGCAGAGCTTGCAGTGAGCCAAGATCGC
GCCACTGCACTCCAGCCTAGGTGACAGGGCGAGACTCCATCTCAAAACCAATAAAAATAA
AAATAAAAATAAATAAATAAAAGTAACTTGGTAAGTTTTAACAGCTTTGATCATAATAAA
ATAGCAGCAAGAGCTCCCAGCACAGGAGCCATAAATGGCCAGCGTATTTCGTAAGTTCGC
[G,T]
TTTGTTCTTTTCAGTGCTTTGCTCTTGTTGTGTATAAGTCAGCTCTTTCTGATGCTGGTT
CAAAACCACAGGCTCCAGAATCCAGTTCCTTCTGTGAACATGACTGTTGGCCTTATGTTG
CTTCAGCAGTTTAAAAGCTCATATTCTTTGTGTCTCTTGACTCGAAGGGAAGATGTTTTG
TAATACTGTTGGAGCCCTCTTGACTAATCATGTGGTCGAGCTGAGGTTGTCCTCTGTCCC
CCCTTTTGTACACGCCACAGCTGAGCTGCTGCTGAGAAGTGTATAACTGCATTTGTTATA

39550 ATGCTGGTTCAAAACCACAGGCTCCAGAATCCAGTTCCTTCTGTGAACATGACTGTTGGC
CTTATGTTGCTTCAGCAGTTTAAAAGCTCATATTCTTTGTGTCTCTTGACTCGAAGGGAA
GATGTTTTGTAATACTGTTGGAGCCCTCTTGACTAATCATGTGGTCGAGCTGAGGTTGTC
CTCTGTCCCCCCTTTTGTACACGCCACAGCTGAGCTGCTGCTGAGAAGTGTATAACTGCA
TTTGTTATACAAATGTCTTCCTTTTTGTCTGGGCTGGGGTCTTTGTGTGTGTGGGGGGGG
[T,G]
GATTAGGGGAGAGTAGGGAGAGGGCTGTTCCTGGCTGGCTGCTTCCTGAGATATCTACCT
TGTTGAGTGTCTCTTCATAGGCACTTTAACTCACAGAAGACATTTAGTGCCAGAAGGGGT
TTTATTTGCCCCACATGTCTGCATAGTCGATTGCTGCTTCTGGAGTTAGTTAAAGTCATT
TTCCATGGTGGCAAAACAGATACCCGTGCTGTTGAACCCTGGGGGCTGCTGATGCTGATT
TGGTTTGGACATCCTTCTCTTCTTCCCACTTTGTGTTAGTGGGAGGCTCGCTCTTCTTGC

42281 GCACAGCTCCAGGGTAAGTGTGAGGCAGGAGGCATGAATCCATTCTTTCCCTGGTGTGTT
AGTCCATTTGCGTTGTTATAAAGAAGCACCTGAGACTGGGTAATTTATAAAGAAAAGAGG
TTTATTTTGGCTCATGGCTCTGCAGGCTGTACAGGAAGTGTGATGCCAGCATCTGCTTCT
GGTGAGGGCCTCAGGAAGCTTCTAATCATGGCAGAAGGCAAAGGGGGAGCAGGCTTTATA
TGGCAAGACAGGGAGCAAGGAGAAGGGAGGTACCAGGCTCTTTTAAACAACAGCTCTCTC
[A,G]
GGGAGGGCCCCAAGTCATTCATGAGGGATTTGCCCCCACGACTCAAACACTTCCCACCAG
GCCCCACCTCTGACATTGGGGATCACATTTCAACATGAAATTTGGAGGGGATCCAAACCA
TATTACCTGGTAAGTCCTTGTTTCCACATGTCTCTCATCTTACTGCAGGGAGTGCTATTC
TCTTTTGTTTGTTTTTATGGCTCCTCAAAAATCAACTTTAGACATTTCAGTTTAAAGTGT
TTCTTAAAAATCTGGTCTCTAAATGCAATCCAATCCTTCAGCTGCTCAGCCAAAGAAGCA

FIGURE 3AA

42321 CATTCTTTCCCTGGTGTGTTAGTCCATTTGCGTTGTTATAAAGAAGCACCTGAGACTGGG
TAATTTATAAAGAAAAGAGGTTTATTTTGGCTCATGGCTCTGCAGGCTGTACAGGAAGTG
TGATGCCAGCATCTGCTTCTGGTGAGGGCCTCAGGAAGCTTCTAATCATGGCAGAAGGCA
AAGGGGGAGCAGGCTTTATATGGCAAGACAGGGAGCAAGGAGAAGGGAGGTACCAGGCTC
TTTTAAACAACAGCTCTCTCAGGGAGGGCCCCAAGTCATTCATGAGGGATTTGCCCCCAC
[G,A]
ACTCAAACACTTCCCACCAGGCCCCACCTCTGACATTGGGGATCACATTTCAACATGAAA
TTTGGAGGGGATCCAAACCATATTACCTGGTAAGTCCTTGTTTCCACATGTCTCTCATCT
TACTGCAGGGAGTGCTATTCTCTTTTGTTTGTTTTTATGGCTCCTCAAAAATCAACTTTA
GACATTTCAGTTTAAAGTGTTTCTTAAAAATCTGGTCTCTAAATGCAATCCAATCCTTCA
GCTGCTCAGCCAAAGAAGCAGTGATCGATGTAGACATTGGCTGCCTTGGACTGAGATGTT

42563 TTAAACAACAGCTCTCTCAGGGAGGGCCCCAAGTCATTCATGAGGGATTTGCCCCCACGA
CTCAAACACTTCCCACCAGGCCCCACCTCTGACATTGGGGATCACATTTCAACATGAAAT
TTGGAGGGGATCCAAACCATATTACCTGGTAAGTCCTTGTTTCCACATGTCTCTCATCTT
ACTGCAGGGAGTGCTATTCTCTTTTGTTTGTTTTTATGGCTCCTCAAAAATCAACTTTAG
ACATTTCAGTTTAAAGTGTTTCTTAAAAATCTGGTCTCTAAATGCAATCCAATCCTTCAG
[G,C]
TGCTCAGCCAAAGAAGCAGTGATCGATGTAGACATTGGCTGCCTTGGACTGAGATGTTCT
GGCAGTCTCACCAGTGTGGTGCCTTCCTTAGAGTGACTTGACTGCATTTTCGCTTTACAG
AATGAACTTAGAAGCAAACCTCTCATATAAAATGTAACCCTCTCGTAGGAATCAATGAGG
TAGTAGATAAGCTCTGGATGTCTGTATCAAGGCTGGGAGCATCCAGCTGTAGCCCAGCAG
TAGGAAAGACAATCTGTCAAACTATATTTGATTGCTAACAGGTTAGTAACTAACAGGAAG

42675 CATGAAATTTGGAGGGGATCCAAACCATATTACCTGGTAAGTCCTTGTTTCCACATGTCT
CTCATCTTACTGCAGGGAGTGCTATTCTCTTTTGTTTGTTTTTATGGCTCCTCAAAAATC
AACTTTAGACATTTCAGTTTAAAGTGTTTCTTAAAAATCTGGTCTCTAAATGCAATCCAA
TCCTTCAGCTGCTCAGCCAAAGAAGCAGTGATCGATGTAGACATTGGCTGCCTTGGACTG
AGATGTTCTGGCAGTCTCACCAGTGTGGTGCCTTCCTTAGAGTGACTTGACTGCATTTTC
[G,A]
CTTTACAGAATGAACTTAGAAGCAAACCTCTCATATAAAATGTAACCCTCTCGTAGGAAT
CAATGAGGTAGTAGATAAGCTCTGGATGTCTGTATCAAGGCTGGGAGCATCCAGCTGTAG
CCCAGCAGTAGGAAAGACAATCTGTCAAACTATATTTGATTGCTAACAGGTTAGTAACTA
ACAGGAAGTCATGCACTGTAGCAGGATGTACTTTTCATGGCCAAAAAGATGAGTACTAAT
GATGATAACATTAACAGGTAAGACATCCCTACTGTACACCAGGCCTTTTGTGAGGCACCT

42908 TGGACTGAGATGTTCTGGCAGTCTCACCAGTGTGGTGCCTTCCTTAGAGTGACTTGACTG
CATTTTCGCTTTACAGAATGAACTTAGAAGCAAACCTCTCATATAAAATGTAACCCTCTC
GTAGGAATCAATGAGGTAGTAGATAAGCTCTGGATGTCTGTATCAAGGCTGGGAGCATCC
AGCTGTAGCCCAGCAGTAGGAAAGACAATCTGTCAAACTATATTTGATTGCTAACAGGTT
AGTAACTAACAGGAAGTCATGCACTGTAGCAGGATGTACTTTTCATGGCCAAAAAGATGA
[G,A]
TACTAATGATGATAACATTAACAGGTAAGACATCCCTACTGTACACCAGGCCTTTTGTGA
GGCACCTGCATAACCTCATTTGACCATCATGACATCTCTATGATTCAGGAGCAGTTAATA
TCCCCATTTTGCCAACAAGAAAACTGGGGAATAGAAAGGTACCATACCTTCCCCAATGTC
ACTCAGCTAATTAGCAGCAGAGCCAGGATCTGAACACAAGAACCTAGTTCCAGAGCCCAC
AGGCCTCAATAAACCTGTGAAACACTGGCCTTTGCCCACCTGGTGGAAAGATCGGTGAGA

43358 AATAGAAAGGTACCATACCTTCCCCAATGTCACTCAGCTAATTAGCAGCAGAGCCAGGAT
CTGAACACAAGAACCTAGTTCCAGAGCCCACAGGCCTCAATAAACCTGTGAAACACTGGC
CTTTGCCCACCTGGTGGAAAGATCGGTGAGATGGGAAGCGTGGGGTCAGTGGGCACTAGG
ATGGGTGTATTCGGTGAAGCCTCCTCCTGCTTACAGCACTGTCTGGCAGTGTTGACAATG
GCTGGTATGGCACGGAAGCCGATGGCACCTCCTGCGGCAGTGCACCATTGGTCTTCGTCA
[-,G]
TTCCTCCTTCCTGGCTCACCCGTGGCTGAGTTTCAGATGTGAGAGCCAGTGGGTGTCCTG
TCACAGAGATACGGTCGTCGTGTGGGGCTTCGCCAGGGGTCAGCCTGCAGATAGAACTGC
TTTTTTTCACCTGTATCAAAATGCTCTGTGAAATGCGGTTTTATCACGGTGTCTTTCCAG
AAGGCGGGGTTTCTTTTCCTATTTGGTTTCTTGTCAGTCAGGTAGAGATGTTTGTGTTGG
AGGCTCCCTGAGTGGTAAGAAAATGAGCAGCTGCTCAGGAACGTCCACCTCCTTTTCTTC

43371 CATACCTTCCCCAATGTCACTCAGCTAATTAGCAGCAGAGCCAGGATCTGAACACAAGAA
CCTAGTTCCAGAGCCCACAGGCCTCAATAAACCTGTGAAACACTGGCCTTTGCCCACCTG
GTGGAAAGATCGGTGAGATGGGAAGCGTGGGGTCAGTGGGCACTAGGATGGGTGTATTCG
GTGAAGCCTCCTCCTGCTTACAGCACTGTCTGGCAGTGTTGACAATGGCTGGTATGGCAC
GGAAGCCGATGGCACCTCCTGCGGCAGTGCACCATTGGTCTTCGTCAGTTCCTCCTTCCT
[G,C]
GCTCACCCGTGGCTGAGTTTCAGATGTGAGAGCCAGTGGGTGTCCTGTCACAGAGATACG

FIGURE 3BB

GTCGTCGTGTGGGGCTTCGCCAGGGGTCAGCCTGCAGATAGAACTGCTTTTTTTCACCTG
TATCAAAATGCTCTGTGAAATGCGGTTTTATCACGGTGTCTTTCCAGAAGGCGGGGTTTC
TTTTCCTATTTGGTTTCTTGTCAGTCAGGTAGAGATGTTTGTGTTGGAGGCTCCCTGAGT
GGTAAGAAAATGAGCAGCTGCTCAGGAACGTCCACCTCCTTTTCTTCTCCCTACCCTCCC

44796 GCCATCGCTCACCTGTACCTATTTACACCCAGAACTTTCCAGCTCCCCCTCATCATGCCT
CCTCCTTCCTACCTGCCTCCCCTCTGCTGGTGCACCTCGCCCAACTCATTCTTACTGCAC
AGTTCACTTTATTTAACAATTTTCATGTCCCCCACCTCATGTTTTCACCTTTTACTGGGC
CAGGCATAGATTAAGTAACTGGGAACGCCCCCTCTTTATAAAGCTGGGCTTCTTTCTCAT
CTCTCTCCCAAATGTTGTATACTCAGTATTCTTCCTATTCGAGTCTCCAGGGGGTGGCTG
[G,A]
ACCTACCTGGTCATTTGAAACAGGCCCCCAAGCTGGAGTTTTTAATCTGGACTCTCTGGC
TTGCTGTGACCCCTAAGGCAATGCTTCTCTTCCCTGGTATTCCTTAGTGTGGGTCACAGT
ACTGTGTTCTTAGTTGCTTTAGCTCTTAAAACATACGAAGTGTTGCCTAAACTGAAAATA
TTTATCTTTTATTTAAAATCAGATTTTTGTTTTTAGACTGTCTTAGATCTGGGGCTATTA
CGAATCACTTCTTCTTCAGTAAACTTTGACTCAACTTCTCCTGCTGAAAAGAAGCTCGCT

45820 GGGAAAAGGGCAGAACCAGTGCCCGGCCCCACACTGCCTCTGTGGCCTGGACTTTGAAAG
GAACCCACTGAACACTAATTATGAGCCCTGTCTTTCCCCCAGAATGCCTCCCTGGGTTTC
ACAAACAGCCTTGAGGTTGGCCCTCCTCAAGGTCAGCCTTCAGATTTGGGAGCAAACTTC
AGAGAAGGCAGAGGAAGATACATTGCCTTGCTGTGGGCTGCCTCTTCTTTCCTCTTGGTG
TGCGAAGTATTTCAGAAGGCCATTGATGAATTCCCCCTCTTTAGCTGTGTATTTGTGCAC
[A,G]
TGTGTGTGTACGTGCGTGTGTGTGTGTTCCTGTGTAAGTAACAGACCAGACTCCTTTT
CTCTTCTGTCCCGTCACCAGGCTCTTGCTTCACTGCAGATACAGTTCACTCTGAAAGCTG
GTTGAAGGAGAGCAGCAAAAATGTATCAGGGGTTTTGCTTCTGTGTTTCGCCAAAGCTCA
TAAGGGCTGTGACCCACCCATATGGCCCCAGTTTTTTCTGTCTCTTCTGTTCCAAAGCCA
GGAGAGCTGACTTCCAGGTGAAGGGATGGGAAAAGTGGACTCTCATTGTAGTGACTCCCA

FIGURE 3CC